



RESEARCH CENTER

FIELD

Activity Report 2015

Section New Results

Edition: 2016-03-21

ALGORITHMICS, PROGRAMMING, SOFTWARE AND ARCHITECTURE

1. ALF Project-Team	9
2. ANTIQUE Project-Team	20
3. AOSTE Project-Team	26
4. ARIC Project-Team	33
5. ATEAMS Project-Team	40
6. CAIRN Project-Team	42
7. CAMUS Team	50
8. CAMEL Project-Team	60
9. CARTE Project-Team	62
10. CASCADE Project-Team	65
11. CASSIS Project-Team	66
12. CELTIQUE Project-Team	76
13. COMETE Project-Team	79
14. COMPSYS Project-Team	83
15. CONVECS Project-Team	90
16. CORSE Team	100
17. CRYPT Team (section vide)	107
18. DECENTRALISE Team	108
19. DEDUCTEAM Team	110
20. DICE Team	113
21. DREAMPAL Project-Team	116
22. ESTASYS Team	122
23. GALAAD2 Team	133
24. GALLIUM Project-Team	137
25. GEOMETRICA Project-Team	145
26. GRACE Project-Team	150
27. HYCOMES Team	155
28. LFANT Project-Team	157
29. MARELLE Project-Team	159
30. MEXICO Project-Team	162
31. MUTANT Project-Team	165
32. PARKAS Project-Team	169
33. PARSIFAL Project-Team	171
34. PI.R2 Project-Team	177
35. POLSYS Project-Team	182
36. POSET Team	189
37. POSTALE Team	191
38. PRIVATICS Project-Team	194
39. PROSECCO Project-Team	198
40. SECRET Project-Team	201

41. SPADES Project-Team	206
42. SPECFUN Project-Team	212
43. SUMO Project-Team	216
44. TASC Project-Team	223
45. TEA Project-Team	225
46. TOCCATA Project-Team	229
47. VEGAS Project-Team	233
48. VERIDIS Project-Team	236

APPLIED MATHEMATICS, COMPUTATION AND SIMULATION

49. ACUMES Team	243
50. APICS Project-Team	248
51. ASPI Project-Team	256
52. BIPOP Project-Team	260
53. CAGIRE Team	265
54. CARDAMOM Team	267
55. COMMANDS Project-Team	274
56. CQFD Project-Team	278
57. DEFI Project-Team	285
58. DISCO Project-Team	290
59. DOLPHIN Project-Team	301
60. ECUADOR Project-Team	306
61. GAMMA3 Project-Team	310
62. GECO Project-Team	315
63. GEOSTAT Project-Team	319
64. I4S Project-Team	328
65. INOCS Team (section vide)	333
66. IPSO Project-Team	334
67. MATHERIALS Project-Team	342
68. MATHRISK Project-Team	350
69. Maxplus Team	353
70. MCTAO Project-Team	371
71. MEMPHIS Team	372
72. MEPHYSTO Team	382
73. MISTIS Project-Team	386
74. MODAL Project-Team	392
75. MOKAPLAN Project-Team	395
76. NACHOS Project-Team	399
77. NANO-D Project-Team	411
78. NECS Project-Team	432
79. NON-A Project-Team	439
80. POEMS Project-Team	447

81. QUANTIC Project-Team	456
82. RAPSODI Team	466
83. REALOPT Project-Team	469
84. SELECT Project-Team	473
85. SEQUEL Project-Team	476
86. SIERRA Project-Team	485
87. SPHINX Team	493
88. TAO Project-Team	496
89. TOSCA Project-Team	500

DIGITAL HEALTH, BIOLOGY AND EARTH

90. ABS Project-Team	505
91. AIRSEA Team	508
92. AMIB Project-Team	517
93. ANGE Project-Team	520
94. ARAMIS Project-Team	524
95. ASCLEPIOS Project-Team	536
96. ATHENA Project-Team	551
97. BEAGLE Project-Team	562
98. BIGS Project-Team	566
99. BIOCORE Project-Team	576
100. BONSAI Project-Team	584
101. CAPSID Project-Team	586
102. CARMEN Team	587
103. CASTOR Project-Team	590
104. CLIME Project-Team	595
105. COFFEE Project-Team (section vide)	601
106. DEMAR Project-Team	602
107. DRACULA Project-Team	614
108. DYLISS Project-Team	618
109. ERABLE Project-Team	622
110. FLUMINANCE Project-Team	631
111. GALEN Project-Team	636
112. GENSCALE Project-Team	640
113. IBIS Project-Team	646
114. LEMON Team	651
115. LIFEWARE Project-Team	655
116. M3DISIM Team	660
117. MAGIQUE-3D Project-Team	666
118. MAMBA Project-Team	672
119. MIMESIS Team	680
120. MNEMOSYNE Project-Team	686

121. MODEMIC Project-Team	688
122. Monc Team	693
123. MORPHEME Project-Team	700
124. MYCENAE Project-Team	709
125. NEUROMATHCOMP Project-Team	715
126. NEUROSYS Project-Team	724
127. NUMED Project-Team (section vide)	728
128. PARIETAL Project-Team	729
129. PLEIADE Team	738
130. POPIX Team	740
131. REO Project-Team	741
132. SAGE Project-Team	746
133. SERENA Team	751
134. SERPICO Project-Team	752
135. SISTM Project-Team	767
136. STEEP Project-Team	768
137. TONUS Team	771
138. VIRTUAL PLANTS Project-Team	775
139. VISAGES Project-Team	794

NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING

140. ALPINES Project-Team	800
141. ASAP Project-Team	804
142. ASCOLA Project-Team	813
143. ATLANMODELS Team	820
144. AVALON Project-Team	823
145. CIDRE Project-Team	831
146. COAST Project-Team	837
147. COATI Project-Team	840
148. CTRL-A Team	850
149. DANTE Project-Team	854
150. DIANA Project-Team	860
151. DIONYSOS Project-Team	864
152. DIVERSE Project-Team	873
153. DYOGENE Project-Team	880
154. EVA Team	890
155. FOCUS Project-Team	898
156. FUN Project-Team	903
157. GANG Project-Team	908
158. HIEPACS Project-Team	917
159. INDES Project-Team	927
160. INFINE Team	931

161. KERDATA Project-Team	937
162. MADYNES Project-Team	943
163. MAESTRO Project-Team	954
164. MESCAL Project-Team	965
165. MIMOVE Team	971
166. MOAIS Project-Team	975
167. MUSE Team	977
168. MYRIADS Project-Team	980
169. PHOENIX Project-Team	988
170. RAP Project-Team	990
171. REGAL Project-Team	994
172. RMOD Project-Team	998
173. ROMA Project-Team	1004
174. SCALE Team	1015
175. SOCRATE Project-Team	1019
176. SPIRALS Project-Team	1027
177. STORM Team	1028
178. TACOMA Team	1032
179. TADAAM Team	1038
180. URBANET Team	1043
181. WHISPER Project-Team	1048

PERCEPTION, COGNITION AND INTERACTION

182. ALICE Project-Team	1051
183. ALPAGE Project-Team	1056
184. AVIZ Project-Team	1061
185. AYIN Team	1068
186. Chroma Team	1079
187. DAHU Project-Team	1094
188. DEFROST Team	1095
189. DREAM Project-Team	1097
190. EX-SITU Team	1103
191. EXMO Project-Team	1109
192. FLOWERS Project-Team	1112
193. GRAPHDECO Project-Team	1127
194. GRAPHIK Project-Team	1139
195. HEPHAISTOS Project-Team	1143
196. HYBRID Project-Team	1147
197. ILDA Team	1159
198. IMAGINE Project-Team	1162
199. LAGADIC Project-Team	1171
200. LARSEN Team	1179

201. LEAR Project-Team	1185
202. LINKMEDIA Project-Team	1198
203. LINKS Team	1206
204. MAGNET Team	1208
205. MAGRIT Project-Team	1210
206. MANAO Project-Team	1214
207. MAVERICK Project-Team	1220
208. MIMETIC Project-Team	1226
209. MINT Project-Team	1234
210. Mjolnir Team	1235
211. MORPHEO Project-Team	1240
212. MULTISPEECH Project-Team	1247
213. OAK Project-Team	1253
214. ORPAILLEUR Project-Team	1255
215. PANAMA Project-Team	1262
216. PERCEPTION Project-Team	1273
217. POTIOC Project-Team	1277
218. PRIMA Project-Team	1287
219. RITS Project-Team	1290
220. SEMAGRAMME Project-Team	1298
221. SIROCCO Project-Team	1301
222. SMIS Project-Team	1309
223. STARS Project-Team	1312
224. TITANE Project-Team	1346
225. TYREX Project-Team	1359
226. WILLOW Project-Team	1362
227. WIMMICS Project-Team	1377
228. ZENITH Project-Team	1385

ALF Project-Team

7. New Results

7.1. Processor Architecture

Participants: Pierre Michaud, Bharath Narasimha Swamy, Sylvain Collange, Erven Rohou, André Seznec, Arthur Perais, Surya Khizakanchery Natarajan, Sajith Kalathingal, Tao Sun, Andrea Mondelli, Aswinkumar Sridharan, Biswabandan Panda, Fernando Endo.

Processor, cache, locality, memory hierarchy, branch prediction, multicore, power, temperature

Multicore processors have now become mainstream for both general-purpose and embedded computing. Instead of working on improving the architecture of the next generation multicore, with the DAL project, we deliberately anticipate the next few generations of multicores. While multicores featuring 1000s of cores might become feasible around 2020, there are strong indications that sequential programming style will continue to be dominant. Even future mainstream parallel applications will exhibit large sequential sections. Amdahl's law indicates that high performance on these sequential sections is needed to enable overall high performance on the whole application. On many (most) applications, the effective performance of future computer systems using a 1000-core processor chip will significantly depend on their performance on both sequential code sections and single threads.

We envision that, around 2020, the processor chips will feature a few complex cores and many (maybe 1000's) simpler, more silicon and power effective cores.

In the DAL research project, <https://team.inria.fr/alf/members/andre-seznec/defying-amdahls-law-dal/>, we explore the microarchitecture techniques that will be needed to enable high performance on such heterogeneous processor chips. Very high performance will be required on both sequential sections, -legacy sequential codes, sequential sections of parallel applications-, and critical threads on parallel applications, -e.g. the main thread controlling the application. Our research focuses essentially on enhancing single process performance.

7.1.1. Microarchitecture

7.1.1.1. Branch prediction

Participant: André Seznec.

This research was done in collaboration with Joshua San Miguel and Jorge Albericio from University of Toronto

The most efficient branch predictors proposed in academic literature exploit both global branch history and local branch history. However, local history branch predictor components introduce major design challenges, particularly for the management of speculative histories. Therefore, most effective hardware designs use only global history components and very limited forms of local histories such as a loop predictor. The wormhole (WH) branch predictor was recently introduced to exploit branch outcome correlation in multidimensional loops. For some branches encapsulated in a multidimensional loop, their outcomes are correlated with those of the same branch in neighbor iterations, but in the previous outer loop iteration. Unfortunately, the practical implementation of the WH predictor is even more challenging than the implementation of local history predictors.

In [36], we introduce practical predictor components to exploit this branch outcome correlation in multidimensional loops: the IMLI-based predictor components. The iteration index of the inner most loop in an application can be efficiently monitored at instruction fetch time using the Inner Most Loop Iteration (IMLI) counter. The outcomes of some branches are strongly correlated with the value of this IMLI counter. A single PC+IMLI counter indexed table, the IMLI-SIC table, added to a neural component of any recent predictor (TAGE-based or perceptron-inspired) captures this correlation. Moreover, using the IMLI counter, one can efficiently manage the very long local histories of branches that are targeted by the WH predictor. A second IMLI-based component, IMLI-OH, allows for tracking the same set of hard-to-predict branches as WH. Managing the speculative states of the IMLI-based predictor components is quite simple. Our experiments show that augmenting a state-of-the-art global history predictor with IMLI components outperforms previous state-of-the-art academic predictors leveraging local and global history at much lower hardware complexity (i.e., smaller storage budget, smaller number of tables and simpler management of speculative states).

7.1.1.2. Revisiting Value Prediction

Participants: Arthur Perais, André Sez nec.

Value prediction was proposed in the mid 90's to enhance the performance of high-end microprocessors. The research on Value Prediction techniques almost vanished in the early 2000's as it was more effective to increase the number of cores than to dedicate some silicon area to Value Prediction. However high end processor chips currently feature 8-16 high-end cores and the technology will allow to implement 50-100 of such cores on a single die in a foreseeable future. Amdahl's law suggests that the performance of most workloads will not scale to that level. Therefore, dedicating more silicon area to value prediction in high-end cores might be considered as worthwhile for future multicores.

At a first step, we showed that all predictors are amenable to very high accuracy at the cost of some loss on prediction coverage [7]. This greatly diminishes the number of value mispredictions and allows to delay validation until commit-time. As such, no complexity is added in the out-of-order engine because of VP (save for ports on the register file) and pipeline squashing at commit-time can be used to recover.

This allows to leverage the possibility of validating predictions at commit to introduce a new microarchitecture, EOLE [19]. EOLE features *Early Execution* to execute simple instructions whose operands are ready in parallel with Rename and *Late Execution* to execute simple predicted instructions and high confidence branches just before Commit. EOLE depends on Value Prediction to provide operands for *Early Execution* and predicted instructions for *Late Execution*. However, Value Prediction requires EOLE to become truly practical. That is, EOLE allows to reduce the out-of-order issue-width by 33% without impeding performance. As such, the number of ports on the register file diminishes. Furthermore, optimizations of the register file such as *banking* further reduce the number of required ports. Overall EOLE possesses a register file whose complexity is on-par with that of a regular wider-issue superscalar while the out-of-order components (scheduler, bypass) are greatly simplified. Moreover, thanks to Value Prediction, speedup is obtained on many benchmarks of the SPEC'00/'06 suite.

However complexity in the value predictor infrastructure itself is also problematic. First, multiple predictions must be generated each cycle, but multi-ported structures should be avoided. Second, the predictor should be small enough to be considered for implementation, yet coverage must remain high enough to increase performance. In [32], to address these remaining concerns, we first propose a block-based value prediction scheme mimicking current instruction fetch mechanisms, BeBoP. It associates the predicted values with a fetch block rather than distinct instructions. Second, to remedy the storage issue, we present the Differential VTAGE predictor. This new tightly coupled hybrid predictor covers instructions predictable by both VTAGE and Stride-based value predictors, and its hardware cost and complexity can be made similar to those of a modern branch predictor. Third, we show that block-based value prediction allows to implement the checkpointing mechanism needed to provide D-VTAGE with last computed/predicted values at moderate cost. Overall, we establish that EOLE with a 32.8KB block-based D-VTAGE predictor and a 4-issue OoO engine can significantly outperform a baseline 6-issue superscalar processor, by up to 62.2 % and 11.2 % on average (gmean), on our benchmark set.

The overall study on value prediction is presented in Arthur Perais's PhD [14].

7.1.1.3. Cost-Effective Speculative Scheduling in High Performance Processors

Participants: André Seznec, Arthur Perais, Pierre Michaud.

This study was done in collaboration with Andreas Sembrant and Erik Hagersten from Upsala University

To maximize performance, out-of-order execution processors sometimes issue instructions without having the guarantee that operands will be available in time; e.g. loads are typically assumed to hit in the L1 cache and dependent instructions are issued assuming a L1 hit. This form of speculation (that we refer to as speculative scheduling) has been used for two decades in real processors, but has received little attention from the research community. In particular, as pipeline depth grows and the distance between the Issue and the Execute stages increases, it becomes critical to issue dependents on variable-latency instructions as soon as possible, rather than to wait for the actual cycle at which the result becomes available. Unfortunately, due to the uncertain nature of speculative scheduling, the scheduler may wrongly issue an instruction that will not have its source(s) on the bypass network when it reaches the Execute stage. Therefore, this instruction must be canceled and replayed, which can potentially impair performance and increase energy consumption.

In [31] we focus on ways to reduce the number of replays that are agnostic of the replay scheme. First, we propose an easily implementable, low-cost solution to reduce the number of replays caused by L1 bank conflicts. Schedule Shifting always assumes that, given a dual-load issue capacity, the second load issued in a given cycle will be delayed because of a bank conflict. Its dependents are thus always issued with a corresponding delay. Second, we also improve on existing L1 hit/miss prediction schemes by taking into account instruction criticality. That is, for some criterion of criticality and for loads whose hit/miss behavior is hard to predict, we show that it is more cost-effective to stall dependents if the load is not predicted critical. In total, in our experiments assuming a 4-cycle issue-to-execute delay, we found that the vast majority of instructions replays due to L1 data cache banks conflicts and L1 hit mispredictions can be avoided, thus leading to a 3.4% performance gain and a 13.4% decrease in the number of issued instructions, over a baseline speculative scheduling scheme.

7.1.1.4. Criticality-aware Resource Allocation in OOO Processors

Participants: André Seznec, Arthur Perais, Pierre Michaud.

This study was done in collaboration with Andreas Sembrant, Erik Hagersten, David Black-Schaffer and Trevor Carlson from Upsala University.

Modern processors employ large structures (IQ, LSQ, register file, etc.) to expose instruction-level parallelism (ILP) and memory-level parallelism (MLP). These resources are typically allocated to instructions in program order. This wastes resources by allocating resources to instructions that are not yet ready to be executed and by eagerly allocating resources to instructions that are not part of the application's critical path. In [35], we explore the possibility of allocating pipeline resources only when needed to expose MLP, and thereby enabling a processor design with significantly smaller structures, without sacrificing performance. First we identify the classes of instructions that should not reserve resources in program order and evaluate the potential performance gains we could achieve by delaying their allocations. We then use this information to "park" such instructions in a simpler, and therefore more efficient, Long Term Parking (LTP) structure. The LTP stores instructions until they are ready to execute, without allocating pipeline resources, and thereby keeps the pipeline available for instructions that can generate further MLP. LTP can accurately and rapidly identify which instructions to park, park them before they execute, wake them when needed to preserve performance, and do so using a simple queue instead of a complex IQ. We show that even a very simple queue-based LTP design allows us to significantly reduce IQ (64 \rightarrow 32) and register file (128 \rightarrow 96) sizes while retaining MLP performance and improving energy efficiency.

7.1.1.5. Efficient Execution on Guarded Instruction Sets

Participant: André Seznec.

ARM ISA based processors are no longer low complexity processors. Nowadays, ARM ISA based processor manufacturers are struggling to implement medium-end to high-end processor cores which implies implementing a state-of-the-art out-of-order execution engine. Unfortunately providing efficient out-of-order execution on legacy ARM codes may be quite challenging due to guarded instructions.

Predicting the guarded instructions addresses the main serialization impact associated with guarded instructions execution and the multiple definition problem. Moreover, guard prediction allows to use a global branch-and-guard history predictor to predict both branches and guards, often improving branch prediction accuracy. Unfortunately such a global branch-and-guard history predictor requires the systematic use of guard predictions. In that case, poor guard prediction accuracy would lead to poor overall performance on some applications.

Building on top of recent advances in branch prediction and confidence estimation, we propose a hybrid branch and guard predictor, combining a global branch history component and global branch-and-guard history component. The potential gain or loss due to the systematic use of guard prediction is dynamically evaluated at run-time. Two computing modes are enabled: systematic guard prediction and high confidence only guard prediction. Our experiments show that on most applications, an overwhelming majority of guarded instructions are predicted. Therefore a relatively inefficient but simple hardware solution can be used to execute the few unpredicted guarded instructions. Significant performance benefits are observed on most applications while applications with poorly predictable guards do not suffer from performance loss [8].

This study was accepted to ACM Transactions on Architecture and Compiler Optimizations (Dec. 2014) and presented at the HIPEAC conference in January 2015.

7.1.1.6. Clustered microarchitecture

Participants: Andrea Mondelli, Pierre Michaud, André Seznec.

In the last 10 years, the clock frequency of high-end superscalar processors did not increase significantly. Performance keeps being increased mainly by integrating more cores on the same chip and by introducing new instruction set extensions. However, this benefits only to some applications and requires rewriting and/or recompiling these applications. A more general way to increase performance is to increase the IPC, the number of instructions executed per cycle.

In [18], we argue that some of the benefits of technology scaling should be used to increase the IPC of future superscalar cores. Starting from microarchitecture parameters similar to recent commercial high-end cores, we show that an effective way to increase the IPC is to increase the issue width. But this must be done without impacting the clock cycle. We propose to combine two known techniques: clustering and register write specialization. The objective of past work on clustered microarchitecture was to allow a higher clock frequency while minimizing the IPC loss. This led researchers to consider narrow-issue clusters. Our objective, instead, is to increase the IPC without impacting the clock cycle, which means wide-issue clusters. We show that, on a wide-issue dual cluster, a very simple steering policy that sends 64 consecutive instructions to the same cluster, the next 64 instructions to the other cluster, and so on, permits tolerating an inter-cluster delay of several cycles. We also propose a method for decreasing the energy cost of sending results of one cluster to the other cluster.

7.1.1.7. Adaptive Intelligent Memory Systems

Participants: André Seznec, Aswinkumar Sridharan.

Multi-core processors employ shared Last Level Caches (LLC). This trend will continue in the future with large multi-core processors (16 cores and beyond) as well. At the same time, the associativity of this LLC tends to remain in the order of sixteen. Consequently, with large multicore processors, the number of cores that share the LLC becomes larger than the associativity of the cache itself. LLC management policies have been extensively studied for small scale multi-cores (4 to 8 cores) and associativity degree in the 16 range. However, the impact of LLC management on large multi-cores is essentially unknown, in particular when the associativity degree is smaller than the number of cores.

In [43], we introduce Adaptive Discrete and deprioritized Application PrioriTization (ADAPT), an LLC management policy addressing the large multi-cores where the LLC associativity degree is smaller than the number of cores. ADAPT builds on the use of the Footprint-number metric. Footprint-number is defined as the number of unique accesses (block addresses) that an application generates to a cache set in an interval of time. We propose a monitoring mechanism that dynamically samples cache sets to estimate the Footprint-number of applications and classifies them into discrete (distinct and more than two) priority buckets. The cache replacement policy leverages this classification and assigns priorities to cache lines of applications during cache replacement operations. Footprint-number is computed periodically to account the dynamic changes in applications behavior. We further find that de-prioritizing certain applications during cache replacement is beneficial to the overall performance. We evaluate our proposal on 16, 20 and 24-core multi-programmed workloads and discuss other aspects in detail.

[43] has been accepted for publication at the IPDPS 2016 conference.

7.1.1.8. Hardware data prefetching

Participant: Pierre Michaud.

Hardware prefetching is an important feature of modern high-performance processors. When an application's working set is too large to fit in on-chip caches, disabling hardware prefetchers may result in severe performance reduction. We propose a new hardware data prefetcher, the Best-Offset (BO) prefetcher. The BO prefetcher is an offset prefetcher using a new method for selecting the best prefetch offset taking into account prefetch timeliness. The hardware required for implementing the BO prefetcher is very simple. The BO prefetcher won the last Data Prefetching Championship [27].

A paper describing and studying the BO prefetcher has been accepted for publication at the HPCA 2016 conference.

7.1.1.9. Prediction-based superpage-friendly TLB designs

Participant: André Seznec.

This research was done in collaboration with Misel-Myrto Papadopoulou, Xin Tong and Andreas Moshovos from University of Toronto

In [30], we demonstrate that a set of commercial and scale-out applications exhibit significant use of superpages and thus suffer from the fixed and small superpage TLB structures of some modern core designs. Other processors better cope with superpages at the expense of using power-hungry and slow fully-associative TLBs. We consider alternate designs that allow all pages to freely share a single, power-efficient and fast set-associative TLB. We propose a prediction-guided multi-grain TLB design that uses a superpage prediction mechanism to avoid multiple lookups in the common case. In addition, we evaluate the previously proposed skewed TLB which builds on principles similar to those used in skewed associative caches. We enhance the original skewed TLB design by using page size prediction to increase its effective associativity. Our prediction-based multi-grain TLB design delivers more hits and is more power efficient than existing alternatives. The predictor uses a 32-byte prediction table indexed by base register values.

7.1.2. Microarchitecture Performance Modeling

7.1.2.1. Symbiotic scheduling on SMT cores and symmetric multicores

Participant: Pierre Michaud.

This research was done in collaboration with Stijn Eyerman and Wouter Rogiest from Ghent University.

When several independent tasks execute concurrently on a simultaneous multithreaded (SMT) core or on a multicore, they share hardware resources. Hence the execution rate of a task is influenced by the other tasks running at the same time. Based on this observation, Snively and Tullsen proposed *symbiotic* scheduling, i.e., the idea that performance can be increased by co-scheduling tasks that do not stress the same shared resources [63]. They claim that, when the number of concurrent tasks exceeds the number of logical cores, symbiotic scheduling increases performance substantially. A more recent study by Eyerman and Eeckhout reached similar conclusions [54].

We have revisited symbiotic scheduling for SMT cores and symmetric multicores [22], and we obtained very modest throughput gains, which seemingly contradicts the above mentioned studies. We analyzed the reasons for this discrepancy and found that previous studies did not measure throughput but average response time. Response time reductions can be magnified by setting the job arrival rate very close to the maximum throughput, which turns a tiny throughput increase into a large response time reduction. Also, the proposed scheduling policies are approximately equivalent to scheduling the shortest jobs first, which mechanically reduces the average response time independently of any symbiosis effect.

We identified three typical situations where symbiotic scheduling yields little to no throughput gain: (1) most of the time is spent executing a single type of job, or (2) jobs' execution rates barely depend on which other jobs are running concurrently, or (3) jobs' execution rates are proportional to the fraction they get of a certain shared resource (e.g., instruction decode bandwidth in an SMT core). In our experiments, most workloads were close to one of the three situations above.

7.1.2.2. *Modeling multi-threaded programs execution time in the many-core era*

Participants: Surya Khizakanchery Natarajan, Bharath Narasimha Swamy, André Seznec.

Estimating the potential performance of parallel applications on the yet-to-be-designed future many cores is very speculative. The simple models proposed by Amdahl's law (fixed input problem size) or Gustafson's law (fixed number of cores) do not completely capture the scaling behaviour of a multi-threaded (MT) application leading to over estimation of performance in the many-core era. On the other hand, modeling many-core by simulation is too slow to study the applications performance. In [28], [13], we propose a more refined but still tractable, high level empirical performance model for multi-threaded applications, the Serial/Parallel Scaling (SPS) Model to study the scalability and performance of application in many-core era. SPS model learns the application behavior on a given architecture and provides realistic estimates of the performance in future many-cores. Considering both input problem size and the number of cores in modeling, SPS model can help in making high level decisions on the design choice of future many-core applications and architecture. We validate the model on the Many-Integrated Cores (MIC) xeon-phi with 240 logical cores.

7.1.2.3. *Optimal cache replacement*

Participant: Pierre Michaud.

This research was done in collaboration with Mun-Kyu Lee, Jeong Seop Sim and DaeHun Nyang from Inha University.

The replacement policy for a cache is the algorithm, implemented in hardware, selecting a block to evict for making room for an incoming block. This research topic has been revitalized in recent years. The MIN replacement policy, which evicts the block referenced furthest in the future, was introduced by Belady [51] and was later shown to be optimal by Mattson et al. [60]. The MIN policy is an offline policy that cannot be implemented in real processors, as it needs the knowledge of future memory accesses. Still, a possible way to improve online replacement policies would be to emulate the MIN policy, trying to use past references to predict future ones. However, the MIN policy is not intuitive, and Mattson et al.'s proof of optimality is quite involved. We believe that new intuition about the MIN policy will help microarchitects improve cache replacement policies. As a first step toward this goal, we produced a new, intuitive proof of optimality of the MIN policy [17].

7.1.3. *Hardware/Software Approaches*

7.1.3.1. *Helper threads*

Participants: Bharath Narasimha Swamy, André Seznec.

Heterogeneous Many Cores (HMC) architectures that mix many simple/small cores with a few complex/large cores are emerging as a design alternative that can provide both fast sequential performance for single threaded workloads and power-efficient execution for throughput oriented parallel workloads. The availability of many small cores in a HMC presents an opportunity to utilize them as low-power helper cores to accelerate memory-intensive sequential programs mapped to a large core. However, the latency overhead of accessing small cores in a loosely coupled system limits their utility as helper cores. Also, it is not clear if small cores can execute helper threads sufficiently in advance to benefit applications running on a larger, much powerful, core.

In [12] we present a hardware/software framework called core-tethering to support efficient helper threading on heterogeneous many-cores. Core-tethering provides a co-processor like interface to the small cores that (a) enables a large core to directly initiate and control helper execution on the helper core and (b) allows efficient transfer of execution context between the cores, thereby reducing the performance overhead of accessing small cores for helper execution. Our evaluation on a set of memory intensive programs chosen from the standard benchmark suites show that, helper threads using moderately sized small cores can significantly accelerate a larger core compared to using a hardware prefetcher alone. We also find that a small core provides a good trade-off against using an equivalent large core to run helper threads in a HMC.

In summary, despite the latency overheads of accessing prefetched cache lines from the shared L3 cache, helper thread based prefetching on small cores looks as a promising way to improve single thread performance on memory intensive workloads in HMC architectures.

This research was partially done in collaboration with Alain Ketterlin from the Inria Camus project-team in Strasbourg.

7.1.3.2. Branch Prediction and Performance of Interpreter

Participants: Erven Rohou, André Seznec, Bharath Narasimha Swamy.

Interpreters have been used in many contexts. They provide portability and ease of development at the expense of performance. The literature of the past decade covers analysis of why interpreters are slow, and many software techniques to improve them. A large proportion of these works focuses on the dispatch loop, and in particular on the implementation of the switch statement: typically an indirect branch instruction. Folklore attributes a significant penalty to this branch, due to its high misprediction rate. In [34], we revisit this assumption, considering state-of-the-art branch predictors and the three most recent Intel processor generations on current interpreters. Using both hardware counters on Haswell, the latest Intel processor generation, and simulation of the ITTAGE predictor [10], we show that the accuracy of indirect branch prediction is no longer critical for interpreters. We further compare the characteristics of these interpreters and analyze why the indirect branch is less important than before.

7.1.3.3. Augmenting superscalar architecture for efficient many-thread parallel execution

Participants: Sylvain Collange, André Seznec, Sajith Kalathingal.

Threads of Single-Program Multiple-Data (SPMD) applications often exhibit very similar control flows, i.e. they execute the same instructions on different data. In [42] we propose the Dynamic Inter-Thread Vectorization Architecture (DITVA) to leverage this implicit Data Level Parallelism on SPMD applications to create dynamic vector instructions at runtime. DITVA extends an in-order SMT processor with SIMD units with an inter-thread vectorization execution mode. In this mode, identical instructions of several threads running in lockstep are aggregated into a single SIMD instruction. DITVA leverages existing SIMD units and maintains binary compatibility with existing CPU architectures. To balance TLP and DLP, threads are statically grouped into fixed-size warps, inside which threads run in lockstep. At instruction fetch time, if the instruction streams of several threads within a warp are synchronized, then DITVA aggregates the instructions of the threads as dynamic vectors. To maximize vectorization opportunities, we use resource sharing arbitration policies that favor thread synchronization within warps. The policies do not require any compiler hints or modified algorithms for the existing SPMD applications and allow to run unmodified CPU binaries. A dynamic vector instruction is executed as a single unit. This allows to execute m identical instructions from m different threads on m parallel execution lanes while activating the I-fetch, the decode, and the overall pipeline control only once.

Our evaluation on the SPMD applications from the PARSEC and SPLASH benchmarks shows that a 4-warp 4-lane 4-issue DITVA architecture with a realistic bank-interleaved cache achieves 44% higher performance than a 4-thread 4-issue SMT architecture with AVX instructions while fetching and issuing 40 % fewer instructions, achieving an overall 22% energy reduction.

7.2. Compiler, vectorization, interpretation

Participants: Erven Rohou, Emmanuel Riou, Bharath Narasimha Swamy, Arjun Suresh, André Seznec, Nabil Hallou, Sylvain Collange.

7.2.1. Improving sequential performance through memoization

Participants: Erven Rohou, Emmanuel Riou, Bharath Narasimha Swamy, André Seznec, Arjun Suresh.

Many applications perform repetitive computations, even when properly programmed and optimized. Performance can be improved by caching results of pure functions, and retrieving them instead of recomputing a result (a technique called memoization).

We propose [20] a simple technique for enabling software memoization of any dynamically linked pure function and we illustrate our framework using a set of computationally expensive pure functions – the transcendental functions.

Our technique does not need the availability of source code and thus can be applied even to commercial applications as well as applications with legacy codes. As far as users are concerned, enabling memoization is as simple as setting an environment variable.

Our framework does not make any specific assumptions about the underlying architecture or compiler tool-chains, and can work with a variety of current architectures.

We present experimental results for x86-64 platform using both gcc and icc compiler tool-chains, and for ARM cortex-A9 platform using gcc. Our experiments include a mix of real world programs and standard benchmark suites: SPEC and Splash2x. On standard benchmark applications that extensively call the transcendental functions we report memoization benefits of upto 16 %, while much higher gains were realized for programs that call the expensive Bessel functions. Memoization was also able to regain a performance loss of 76 % in *bwaves* due to a known performance bug in the gcc libm implementation of *pow* function.

This work has been published in ACM TACO 2015 [20] and accepted for presentation at the International Conference HiPEAC 2016.

7.2.2. Code Obfuscation

Participant: Erven Rohou.

This research is done in collaboration with the group of Prof. Ahmed El-Mahdy at E-JUST, Alexandria, Egypt.

We propose [24] to leverage JIT compilation to make software tamper-proof. The idea is to constantly generate different versions of an application, even while it runs, to make reverse engineering hopeless. More precisely a JIT engine is used to generate new versions of a function each time it is invoked, applying different optimizations, heuristics and parameters to generate diverse binary code. A strong random number generator will guarantee that generated code is not reproducible, though the functionality is the same.

This work was presented in January 2015 at the International Workshop on Dynamic Compilation Everywhere (DCE-2015) [24].

7.2.3. Dynamic Binary Re-vectorization

Participants: Erven Rohou, Nabil Hallou, Emmanuel Riou.

This work is done in collaboration with Philippe Clauss and Alain Ketterlin (Inria CAMUS).

Applications are often under-optimized for the hardware on which they run. Several reasons contribute to this unsatisfying situation, including the use of legacy code, commercial code distributed in binary form, or deployment on compute farms. In fact, backward compatibility of instruction sets guarantees only the functionality, not the best exploitation of the hardware. In particular SIMD instruction sets are always evolving.

We proposed [23] a runtime re-vectorization platform that dynamically adapts applications to execution hardware. The platform is built on top of Padrone. Programs distributed in binary forms are re-vectorized at runtime for the underlying execution hardware. Focusing on the x86 SIMD extensions, we are able to automatically convert loops vectorized for SSE into the more recent and powerful AVX. A lightweight mechanism leverages the sophisticated technology put in a static vectorizer and adjusts, at minimal cost, the width of vectorized loops. We achieve speedups in line with a native compiler targeting AVX. Our re-vectorizer is implemented inside a dynamic optimization platform; its usage is completely transparent to the user and requires neither access to source code nor rewriting binaries.

7.2.4. *Dynamic Parallelization of Binary Executables*

Participants: Erven Rohou, Nabil Hallou, Emmanuel Riou.

We address runtime automatic parallelization of binary executables, assuming no previous knowledge on the executable code. The Padrone platform is used to identify candidate functions and loops. Then we disassemble the loops and convert them to the intermediate representation of the LLVM compiler (thanks to the external tool McSema). This allows us to leverage the power of the polyhedral model for auto-parallelizing loops. Once optimized, new native code is generated just-in-time in the address space of the target process.

Our approach enables user transparent auto-parallelization of legacy and/or commercial applications with auto-parallelization.

This work is done in collaboration with Philippe Clauss (Inria CAMUS).

7.2.5. *Hardware Accelerated JIT Compilation for Embedded VLIW Processors*

Participant: Erven Rohou.

Just-in-time (JIT) compilation is widely used in current embedded systems (mainly because of Java Virtual Machine). When targeting Very Long Instruction Word (VLIW) processors, JIT compilation back-ends grow more complex because of the instruction scheduling phase. This tends to reduce the benefits of JIT compilation for such systems. We propose a hybrid JIT compiler where JIT management is handled in software and the back-end is performed by specialized hardware. Experimental studies show that this approach leads to a compilation up to 15 times faster and 18 times more energy efficient than a pure software compilation.

This work is done in collaboration with the CAIRN team (Steven Derrien and Simon Rokicki).

7.2.6. *Performance Assessment of Sequential Code*

Participant: Erven Rohou.

The advent of multicore and manycore processors, including GPUs, in the customer market encouraged developers to focus on extraction of parallelism. While it is certainly true that parallelism can deliver performance boosts, parallelization is also a very complex and error-prone task, and many applications are still dominated by sequential sections. Micro-architectures have become extremely complex, and they usually do a very good job at executing fast a given sequence of instructions. When they occasionally fail, however, the penalty is severe. Pathological behaviors often have their roots in very low-level details of the micro-architecture, hardly available to the programmer. In [33], we argue that the impact of these low-level features on performance has been overlooked, often relegated to experts. We show that a few metrics can be easily defined to help assess the overall performance of an application, and quickly diagnose a problem. Finally, we illustrate our claim with a simple prototype, along with use cases.

7.2.7. *Compilers for emerging throughput architectures*

Participant: Sylvain Collange.

This work is done in collaboration with Douglas de Couto and Fernando Pereira from UFMG.

The increasing popularity of Graphics Processing Units (GPUs) has brought renewed attention to old problems related to the Single Instruction, Multiple Data execution model. One of these problems is the reconvergence of divergent threads. A divergence happens at a conditional branch when different threads disagree on the path to follow upon reaching this split point. Divergences may impose a heavy burden on the performance of parallel programs.

We have proposed a compiler-level optimization to mitigate the performance loss due to branch divergence on GPUs. This optimization consists in merging function call sites located at different paths that sprout from the same branch. We show that our optimization adds negligible overhead on the compiler. When not applicable, it does not slow down programs and it accelerates substantially those in which it is applicable. As an example, we have been able to speed up the well known SPLASH Fast Fourier Transform benchmark by 11 %.

7.2.8. Deterministic floating-point primitives for high-performance computing

Participant: Sylvain Collange.

This work is done in collaboration with David Defour (UPVD), Stef Graillat and Roman Iakymchuk (LIP6).

Parallel algorithms such as reduction are ubiquitous in parallel programming, and especially high-performance computing. Although these algorithms rely on associativity, they are used on floating-point data, on which operations are not associative. As a result, computations become non-deterministic, and the result may change according to static and dynamic parameters such as machine configuration or task scheduling.

We introduced a solution to compute deterministic sums of floating-point numbers efficiently and with the best possible accuracy. A multi-level algorithm incorporating a filtering stage that uses fast vectorized floating-point expansions and an accumulation stage based on superaccumulators in a high-radix carry-save representation guarantees accuracy to the last bit even on degenerate cases while maintaining high performance in the common cases [16]. Leveraging these algorithms, we build a reproducible BLAS library [49] and extend the approach to triangular solvers [25].

7.3. WCET estimation and optimization

Participants: Hanbing Li, Isabelle Puaut, Erven Rohou, Damien Hardy, Viet Anh Nguyen, Benjamin Rouxel.

7.3.1. WCET estimation for architectures with faulty caches

Participants: Damien Hardy, Isabelle Puaut.

This is joint work with Yannakis Sazeides from University of Cyprus

Fine-grained disabling and reconfiguration of hardware elements (functional units, cache blocks) will become economically necessary to recover from permanent failures, whose rate is expected to increase dramatically in the near future. This fine-grained disabling will lead to degraded performance as compared to a fault-free execution.

Until recently, all static worst-case execution time (WCET) estimation methods were assuming fault-free processors, resulting in unsafe estimates in the presence of faults. The first static WCET estimation technique dealing with the presence of permanent faults in instruction caches was proposed in [4]. This study probabilistically quantified the impact of permanent faults on WCET estimates. It demonstrated that the probabilistic WCET (pWCET) estimates of tasks increase rapidly with the probability of faults as compared to fault-free WCET estimates.

New results show that very simple reliability mechanisms allow mitigating the impact of faulty cache blocks on pWCETs. Two mechanisms, that make part of the cache resilient to faults are analyzed. Experiments show that the gain in pWCET for these two mechanisms are on average 48% and 40% as compared to an architecture with no reliability mechanism.

This work will appear at DATE 2016.

7.3.2. Speeding up Static Probabilistic Timing Analysis

Participants: Damien Hardy, Isabelle Puaut.

This is joint work with Suzana Milutinovic, Jaume Abella, Eduardo Quinones and Francisco J. Cazorla from Barcelona Supercomputing Center.

Probabilistic Timing Analysis (PTA) has emerged recently to derive trustworthy and tight WCET estimates. For its static variant, called SPTA, we identify one of the main elements that jeopardizes its scalability to real-size programs: its high computation time cost. This SPTA's high computational costs are due to convolution, a mathematical operator used by SPTA and also deployed in many domains including signal and image processing.

In [40], we show how convolution is applied in SPTA, and qualitatively and quantitatively evaluate optimizations developed in other domains to reduce convolution time cost when applied to SPTA, and SPTA-specific optimizations. We show that SPTA-specific optimizations provide larger execution time reductions than generic cores.

7.3.3. Traceability of flow information for WCET estimation

Participants: Hanbing Li, Isabelle Puaut, Erven Rohou.

This research is part of the ANR W-SEPT project.

Control-flow information is mandatory for WCET estimation, to guarantee that programs terminate (e.g. provision of bounds for the number of loop iterations) but also to obtain tight estimates (e.g. identification of infeasible or mutually exclusive paths). Such flow information is expressed through annotations, that may be calculated automatically by program/model analysis, or provided manually.

The objective of this work is to address the challenging issue of the mapping and transformation of the flow information from high level down to machine code. In our recent work, we have proposed a framework to systematically transform flow information from source code to machine code. The framework [11] defines a set of formulas to transform flow information for standard compiler optimizations. Transforming the flow information is done within the compiler, in parallel with transforming the code. There thus is no guessing what flow information have become, it is transformed along with the code.

Our most recent results in this framework were to add support for vectorization [26]. We implemented our approach in the LLVM compiler. In addition, we show through measurements on single-path programs that vectorization improves not only average-case performance but also WCETs. The WCET improvement ratio ranges from 1.18x to 1.41x depending on the target architecture on a benchmark suite designed for vectorizing compilers (TSVC).

This work is part of a more general traceability framework, designed and implemented within the ANR W-SEPT project and described in paper [21]. In this paper, we introduce a complete semantic-aware WCET estimation workflow. We introduce some program analysis to find infeasible paths: they can be performed at design, C or binary level, and may take into account information provided by the user. We design an annotation-aware compilation process that enables to trace the infeasible path properties through the program transformations performed by the compilers. Finally, we adapt the WCET estimation tool to take into account the kind of annotations produced by the workflow.

7.3.4. WCET estimation for many core processors

Participants: Viet Anh Nguyen, Damien Hardy, Isabelle Puaut.

This research is part of the PIA Capacités project.

The overall goal of this research is to defined WCET estimation methods for parallel applications running on many-core architectures, such as the Kalray MPPA machine.

Some approaches to reach this goal have been proposed, but they assume the mapping of parallel applications on cores already done. Unfortunately, on architectures with caches, task mapping requires a priori known WCETs for tasks, which in turn requires knowing task mapping (i.e., co-located tasks, co-running tasks) to have tight WCET bounds. Therefore, scheduling parallel applications and estimating their WCET introduce a chicken and egg situation.

In [41], we address this issue by developing an optimal integer linear programming formulation for solving the scheduling problem, whose objective is to minimize the WCET of a parallel application. Our proposed static partitioned non-preemptive mapping strategy addresses the effect of local caches to tighten the estimated WCET of the parallel application. We report preliminary results obtained on synthetic parallel applications.

ANTIQUÉ Project-Team

6. New Results

6.1. Memory Abstraction

6.1.1. Abstraction of arrays based on non contiguous partitions

Participants: Jiangchao Liu, Xavier Rival [correspondant].

Abstract interpretation, Memory abstraction, Array abstract domains. In [19], we studied array abstractions.

Array partitioning analyses split arrays into contiguous partitions to infer properties of cell sets. Such analyses cannot group together non contiguous cells, even when they have similar properties. We proposed an abstract domain which utilizes semantic properties to split array cells into groups. Cells with similar properties will be packed into groups and abstracted together. Additionally, groups are not necessarily contiguous. This abstract domain allows to infer complex array invariants in a fully automatic way. Experiments on examples from the Minix 1.1 memory management demonstrated its effectiveness.

6.1.2. Static analysis for unstructured sharing

Participants: Huisong Li, Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival [correspondant].

Abstract interpretation, Memory abstraction, Separation logic. In [18], we studied the abstraction of shared data-structures.

Shape analysis aims to infer precise structural properties of imperative memory states and has been applied heavily to verify safety properties on imperative code over pointer-based data structures. Recent advances in shape analysis based on separation logic has leveraged summarization predicates that describe unbounded heap regions like lists or trees using inductive definitions. Unfortunately, data structures with *unstructured sharing*, such as graphs, are challenging to describe and reason about in such frameworks. In particular, when the sharing is unstructured, it cannot be described inductively in a local manner. In this work, we proposed a global abstraction of sharing based on set-valued variables that when integrated with inductive definitions enables the specification and shape analysis of structures with unstructured sharing.

6.1.3. Synthesizing short-circuiting validation of data structure invariants

Participants: Yi-Fan Tsai, Devin Coughlin, Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival [correspondant].

In [28], we studied the synthesis of short-circuiting validators for data-structure invariants.

This work introduces *incremental verification-validation*, a novel approach for checking rich data structure invariants expressed as separation logic assertions. Incremental verification-validation combines static verification of separation properties with efficient, *short-circuiting* dynamic validation of arbitrarily rich data constraints. A data structure invariant checker is an inductive predicate in separation logic with an executable interpretation; a short-circuiting checker is an invariant checker that stops checking whenever it detects at *run time* that an assertion for some sub-structure has been fully proven *statically*. At a high level, our approach does two things: it statically proves the separation properties of data structure invariants using a static shape analysis in a standard way but then leverages this proof in a novel manner to synthesize short-circuiting dynamic validation of the data properties. As a consequence, this approach enables dynamic validation to make up for imprecision in sound static analysis while simultaneously leveraging the static verification to make the remaining dynamic validation efficient. This work has shown empirically that short-circuiting can yield asymptotic improvements in dynamic validation, with low overhead over no validation, even in cases where static verification is incomplete.

6.2. Abstract domains

6.2.1. *Abstract domains and solvers for set reasoning*

Participants: Arlen Cox, Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Huisong Li, Xavier Rival [correspondant].

In [15], we studied the abstraction and inference of set properties.

When constructing complex program analyses, it is often useful to reason about not just individual values, but collections of values. Symbolic set abstractions provide building blocks that can be used to partition elements, relate partitions to other partitions, and determine the provenance of multiple values, all without knowing any concrete values. To address the simultaneous challenges of scalability and precision, we formalized and implemented an interface for symbolic set abstractions and constructed multiple abstract domains relying on both specialized data structures and off-the-shelf theorem provers. We developed techniques for lifting existing domains to improve performance and precision. We evaluated these domains on real-world data structure analysis problems.

6.2.2. *Abstraction of optional numerical values*

Participants: Jiangchao Liu, Xavier Rival [correspondant].

In [20], we designed a functor to lift a numerical abstract domain into an abstract domain that accounts for *optional* numerical values.

We proposed a technique to describe properties of numerical stores with optional values, that is, where some variables may have no value. Properties of interest include numerical equalities and inequalities. Our approach lifts common linear inequality based numerical abstract domains into abstract domains describing stores with optional values. This abstraction can be used in order to analyze languages with some form of *option* scalar type. It can also be applied to the construction of abstract domains to describe complex memory properties that introduce symbolic variables, e.g., in order to summarize unbounded sets of program variables, and where these symbolic variables may be undefined, as in some array or shape analyses. We described the general form of abstract states, and propose sound and automatic static analysis algorithms. We evaluated our construction in the case of an array abstract domain.

6.3. Static analysis of JavaScript applications

6.3.1. *Desynchronized multi-state abstractions for open programs in dynamic languages*

Participants: Arlen Cox [correspondant], Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival.

Abstract interpretation, Dynamically typed languages, Verification In [16], we have studied desynchronized multi-state abstractions for open programs in dynamic languages (libraries).

Dynamic language library developers face a challenging problem: ensuring that their libraries will behave correctly for a wide variety of client programs without having access to those client programs. This problem stems from the common use of two defining features for dynamic languages: callbacks into client code and complex manipulation of attribute names within objects. To remedy this problem, we introduced two state-spanning abstractions. To analyze callbacks, the first abstraction desynchronizes a heap, allowing partitions of the heap that may be affected by a callback to an unknown function to be frozen in the state prior to the call. To analyze object attribute manipulation, building upon an abstraction for dynamic language heaps, the second abstraction tracks attribute name/value pairs across the execution of a library. We implemented these abstractions and use them to verify modular specifications of class-, trait-, and mixin-implementing libraries.

6.4. Static analysis of spreadsheet applications

Participants: Tie Cheng [correspondant], Xavier Rival.

Abstract interpretation, Spreadsheet applications, Verification In [14], we have proposed a static analysis to detect type unsafe operations in spreadsheet applications including formulas and macros.

Spreadsheets are widely used, yet are error-prone: they use a weak type system, allowing certain operations that will silently return unexpected results, like comparisons of integer values with string values. However, discovering these issues is hard, since data and formulas can be dynamically set, read or modified. We defined a static analysis that detects all run-time type-unsafe operations in spreadsheets. It is based on an abstract interpretation of spreadsheet applications, including spreadsheet tables, global re-evaluation and associated programs. Our implementation supports the features commonly found in real-world spreadsheets. We ran our analyzer on the EUSES Spreadsheet Corpus. This evaluation shows that our tool is able to automatically verify a large number of real spreadsheets, runs in a reasonable time and discovers complex bugs that are difficult to detect by code review or by testing.

6.5. Distributed systems verification and programming language

Participants: Cezara Drăgoi [correspondant], Thomas Henzinger [IST Austria, Austria], Damien Zufferey [MIT, CSAIL, USA].

Fault-tolerant distributed systems, Programming languages, Verification Fault-tolerant distributed algorithms play an important role in many critical/high-availability applications. These algorithms are notoriously difficult to implement correctly, due to asynchronous communication and the occurrence of faults, such as the network dropping messages or computers crashing. Noteworthy is the lack of automated verification techniques for distributed systems, highly contrasting the mass distribution and development of distributed software. Therefore, our main motivation is to increase the confidence we have in distributed systems using formal verification methods. However, due to the complexity distributed systems have reached, we believe it is no longer realistic nor efficient to assume that high level specifications can be proved when development and verification are two disconnected steps in the software production process. We think that the difficulty does not only come from the algorithms but from the way we think about distributed systems. Therefore, we are interested in finding an appropriate programming model for fault-tolerant distributed algorithms, that increases the confidence we have distributed software. We introduced PSYNC, a domain specific language based on the Heard-Of model, which views asynchronous faulty systems as synchronous ones with an adversarial environment that simulates asynchrony and faults by dropping messages. We defined a runtime system for PSYNC that efficiently executes on asynchronous networks. We formalize the relation between the runtime system and PSYNC in terms of observational refinement. PSYNC introduces a high-level lockstep abstraction (on top of the standard asynchronous semantics), which simplifies the design and implementation of fault-tolerant distributed algorithms and enables automated formal verification. We have implemented an embedding of PSYNC in the SCALA programming language with a runtime system for asynchronous networks. We showed the applicability of PSYNC by implementing several important fault-tolerant distributed algorithms and we compared the implementation of consensus algorithms in PSYNC against implementations in other languages in terms of code size, runtime efficiency, and verification.

6.6. Derivation of Qualitative Dynamical Models from Biochemical Networks

Participants: Wassim Abou-Jaoudé [IBENS], Jérôme Feret [correspondant], Denis Thieffry [IBENS].

Systems biology, Logical models, Automatic derivation As technological advances allow a better identification of cellular networks, more and more molecular data are produced allowing the construction of detailed molecular interaction maps. One strategy to get insights into the dynamical properties of such systems is to derive compact dynamical models from these maps, in order to ease the analysis of their dynamics.

Starting from a case study, we present in [13] a methodology for the derivation of qualitative dynamical models from biochemical networks. Properties are formalized using abstract interpretation. We first abstract states and traces by quotienting the number of instances of chemical species by intervals. Since this abstraction is too coarse to reproduce the properties of interest, we refine it by introducing additional constraints. The resulting abstraction is able to identify the dynamical properties of interest in our case study.

6.7. Annotation of rule-based models with formal semantics to enable creation, analysis, reuse and visualization

Participants: G. Misirli, M. Cavaliere, W. Waites, M. Pocock, C. Madsen, O. Gifellon, R. Honorato-Zimmer, P. Zuliani, V. Danos [correspondant], A. Wipat.

In [35] We present an annotation framework and guidelines for annotating rule-based models, encoded in the commonly used Kappa and BioNetGen languages. Biological systems are complex and challenging to model and therefore model reuse is highly desirable. To promote model reuse, models should include both information about the specifics of simulations and the underlying biology in the form of metadata. The availability of computationally tractable metadata is especially important for the effective automated interpretation and processing of models. Metadata are typically represented as machine-readable annotations which enhance programmatic access to information about models. Rule-based languages have emerged as a modelling framework to represent the complexity of biological systems. Annotation approaches have been widely used for reaction-based formalisms such as SBML. However, rule-based languages still lack a rich annotation framework to add semantic information, such as machine-readable descriptions, to the components of a model. We introduced an annotation framework and guidelines for annotating rule-based models, encoded in the commonly used Kappa and BioNetGen languages. We adapted widely adopted annotation approaches to rule-based models. We initially proposed a syntax to store machine-readable annotations and describe a mapping between rule-based modelling entities, such as agents and rules, and their annotations. We then described an ontology to both annotate these models and capture the information contained therein, and demonstrate annotating these models using examples. Finally, we presented a proof of concept tool for extracting annotations from a model that can be queried and analyzed in a uniform way. The uniform representation of the annotations can be used to facilitate the creation, analysis, reuse and visualization of rule-based models. Although examples are given, using specific implementations the proposed techniques can be applied to rule-based models in general.

6.8. Quantitative genomic analysis of RecA protein binding during DNA double-strand break repair reveals RecBCD action in vivo

Participants: Charlotte Cockram, Milana Filatenkova, Vincent Danos [correspondant], Meriem Karoui, Leach David.

Understanding molecular mechanisms in the context of living cells requires the development of new methods of in vivo biochemical analysis to complement established in vitro biochemistry. A critically important molecular mechanism is genetic recombination, required for the beneficial reassortment of genetic information and for DNA double-strand break repair (DSBR). Central to recombination is the RecA (Rad51) protein that assembles into a spiral filament on DNA and mediates genetic exchange. Here we developed a method that combines chromatin immunoprecipitation with next-generation sequencing (ChIP-Seq) and mathematical modeling to quantify RecA protein binding during the active repair of a single DSB in the chromosome of *Escherichia coli*. In [29] we have used quantitative genomic analysis to infer the key in vivo molecular parameters governing RecA loading by the helicase/ nuclease RecBCD at recombination hot-spots, known as Chi. Our genomic analysis has also revealed that DSBR at the lacZ locus causes a second RecBCD-mediated DSBR event to occur in the ter- minus region of the chromosome, over 1 Mb away.

6.9. Moment Semantics for Reversible Rule-Based Systems

Participants: Vincent Danos [correspondant], Tobias Hinder, Ricardo Honorato-Zimmer, Sandro Stuck.

In [34] we developed a notion of stochastic rewriting over marked graphs – i.e. directed multigraphs with degree constraints. The approach is based on double-pushout (DPO) graph rewriting. Marked graphs are expressive enough to internalize the ‘no-dangling-edge’ condition inherent in DPO rewriting. Our main result is that the linear span of marked graph occurrence-counting functions – or motif functions – form an algebra which is closed under the infinitesimal generator of (the Markov chain associated with) any such rewriting

system. This gives a general procedure to derive the moment semantics of any such rewriting system, as a countable (and recursively enumerable) system of differential equations indexed by motif functions. The differential system describes the time evolution of moments (of any order) of these motif functions under the rewriting system. We illustrate the semantics using the example of preferential attachment networks; a well-studied complex system, which meshes well with our notion of marked graph rewriting. We show how in this case our procedure obtains a finite description of all moments of degree counts for a fixed degree.

6.10. Dirichlet is Natural

Participants: Vincent Danos [correspondant], Ilias Garnier.

In [32] the authors reconstruct a family of higher-order probabilities known as the Dirichlet process.

Giry and Lawvere's categorical treatment of probabilities, based on the probabilistic monad G , offer an elegant and hitherto unexploited treatment of higher-order probabilities. The goal of this paper is to follow this formulation to reconstruct a family of higher-order probabilities known as the Dirichlet process. This family is widely used in non-parametric Bayesian learning.

Given a Polish space X , we build a family of higher-order probabilities in $G(G(X))$ indexed by $M(X)$, the set of non-zero finite measures over X . The construction relies on two ingredients. First, we develop a method to map a zero-dimensional Polish space X to a projective system of finite approximations, the limit of which is a zero-dimensional compactification of X . Second, we use a functorial version of Bochner's probability extension theorem adapted to Polish spaces, where consistent systems of probabilities over a projective system give rise to an actual probability on the limit. These ingredients are combined with known combinatorial properties of Dirichlet processes on finite spaces to obtain the Dirichlet family on X . We prove that the Dirichlet family is a natural transformation from the monad M to GG over Polish spaces, which in particular is continuous in its parameters. This is an improvement on extant constructions of Dirichlet.

6.11. Mechanistic links between cellular trade-offs, gene expression, and growth

Participants: Andrea Weisse, Diego Oyarzun, Vincent Danos [correspondant], Peter Swain.

Intracellular processes rarely work in isolation but continually interact with the rest of the cell. In microbes, for example, we now know that gene expression across the whole genome typically changes with growth rate. The mechanisms driving such global regulation, however, are not well understood. In [36] we considered three trade-offs that, because of limitations in levels of cellular energy, free ribosomes, and proteins, are faced by all living cells and we construct a mechanistic model that comprises these trade-offs. Our model couples gene expression with growth rate and growth rate with a growing population of cells. We show that the model recovers Monod's law for the growth of microbes and two other empirical relationships connecting growth rate to the mass fraction of ribosomes. Further, we can explain growth-related effects in dosage compensation by paralogs and predict host-circuit interactions in synthetic biology. Simulating competitions between strains, we find that the regulation of metabolic pathways may have evolved not to match expression of enzymes to levels of extracellular substrates in changing environments but rather to balance a trade-off between exploiting one type of nutrient over another. Although coarse-grained, the trade-offs that the model embodies are fundamental, and, as such, our modeling framework has potentially wide application, including in both biotechnology and medicine.

6.12. Thermodynamic graph-rewriting

Participants: Vincent Danos [correspondant], Russell Harmer, Ricardo Honorato-Zimmer.

In [33] we developed a new thermodynamic approach to stochastic graph-rewriting. The ingredients are a finite set of reversible graph-rewriting rules called generating rules, a finite set of connected graphs P called energy patterns and an energy cost function. The idea is that the generators define the qualitative dynamics, by showing which transformations are possible, while the energy patterns and cost function specify the long-term probability π of any reachable graph. Given the generators and energy patterns, we construct a finite set of rules which (i) has the same qualitative transition system as the generators; and (ii) when equipped with suitable rates, defines a continuous-time Markov chain of which π is the unique fixed point. The construction relies on the use of site graphs and a technique of ‘growth policy’ for quantitative rule refinement which is of independent interest. This division of labour between the qualitative and long-term quantitative aspects of the dynamics leads to intuitive and concise descriptions for realistic models (see the examples in S4 and S5). It also guarantees thermodynamical consistency (AKA detailed balance), otherwise known to be undecidable, which is important for some applications. Finally, it leads to parsimonious parameterizations of models, again an important point in some applications.

6.13. Kappa Rule-Based Modelling in Synthetic Biology

Participants: John Wilson-Kanamori, Vincent Danos [correspondant], Ty Thomson, Ricardo Honorato-Zimmer.

This [37] is a chapter of a book that provides complete coverage of the computational approaches currently used in Synthetic Biology. Rule-based modeling, an alternative to traditional reaction-based modeling, allows us to intuitively specify biological interactions while abstracting from the underlying combinatorial complexity. One such rule-based modeling formalism is Kappa, which we introduce to readers in this chapter. We discuss the application of Kappa to three modeling scenarios in synthetic biology: a unidirectional switch based on nitrosylase induction in *Saccharomyces cerevisiae*, the repressilator in *Escherichia coli* formed from BioBrick parts, and a light-mediated extension to said repressilator developed by the University of Edinburgh team during iGEM 2010. The second and third scenarios in particular form a case-based introduction to the Kappa BioBrick Framework, allowing us to systematically address the modeling of devices and circuits based on BioBrick parts in Kappa. Through the use of these examples, we highlight the ease with which Kappa can model biological interactions both at the genetic and the protein–protein interaction level, resulting in detailed stochastic models accounting naturally for transcriptional and translational resource usage. We also hope to impart the intuitively modular nature of the modeling processes involved, supported by the introduction of visual representations of Kappa models. Concluding, we explore future endeavors aimed at making modeling of synthetic biology more user-friendly and accessible, taking advantage of the strengths of rule-based modeling in Kappa.

This Chapters focus on computational methods and algorithms for the design of bio-components, insight on CAD programs, analysis techniques, and distributed systems. Written in the highly successful Methods in Molecular Biology series format, the chapters include the kind of detailed description and implementation advice that is crucial for getting optimal results in the laboratory.

Authoritative and practical, Computational Methods in Synthetic Biology serves as a guide to plan *in silico* the *in vivo* or *in vitro* construction of a variety of synthetic bio-circuits.

AOSTE Project-Team

7. New Results

7.1. CCSL as a Logical Clock Calculus Algebra: expressiveness and decidability results

Participants: Robert de Simone, Julien Deantoni, Frédéric Mallet, Qingguo Xu.

CCSL is a language dedicated to the expression of time constraints, based on so-called logical clocks. Its declarative nature is akin to the Lustre or (even closer to) the Signal language, but without values (to clock/event occurrences) and with both synchronous and asynchronous constraints. Solving a set of CCSL constraints amounts to the production of a feasible schedule of the system. While the TimeSquare tool may attempt to generate such a schedule trace by insightful simulation, it is not guaranteed to be complete in its search. So the issue of expressiveness and decidability was left open to this day.

Still, in previous years, we had established the CCSL constraints could be translated into parallel products (extended, transition-labelled) Büchi machines, but some of these machines had to contain integer shift counters, and were thus not fully FSMs. Our (misled) conjecture that CCSL had semilinear, Presburger-arithmetic power was defeated by a new translation expressing (unitary then general) Petri Nets and Vector Addition Systems into CCSL by encoding. The new conjecture that CCSL was then as powerful as Petri Nets was again defeated by a construction interpreting the features of *inhibitor arcs* in CCSL. As such inhibitor arcs extend the expressive power of Petri Nets to become universal (Turing-complete), CCSL enjoys the same universal property (which makes it unfortunately impossible to solve automatically in general).

Despite this negative result we could show that, under natural restrictions such as the assumption that "input" clocks have bounded jitter around a mean rate, and even if those bounds are not exactly known (but may be used as a parameter), then expressiveness remains in the semi-linear, Presburger-arithmetic range.

As a side-effect of this work we provided the translation of CCSL constraints into Büchi components by using a well-defined fragment of the Esterel syntax to express the Buchi automata.

Preliminary results are exposed in a research report. A much more ambitious article is in preparation.

As part of Professor Xu sabbatical in Aoste, we also considered the topic of machine-assisted proof of schedulability using theorem-provers (in our case PVS) [54]).

7.2. Industrial design flow for Embedded System Engineering

Participants: Julien Deantoni, Frédéric Mallet, Marie Agnes Peraldi Frati, Robert de Simone, Ales Mishchenko.

As part of the PIA LEOC Clarity collaborative project we attempt to instill some of our theoretical and methodological ideas into the framework of the (open-source, Polarsys Eclipse) Capella environment. This environment was initially developed inside Thales, under the name ARCADIA/Melody, as a modeling tool flow for System-Level Design in-the-large. As such, several aspects were not fully considered, specially those regarding safe sound simulation semantics at this level, or the role of states and modes in variability regarding both the software applicative and hardware architectural platform models. This research is in part motivated by concrete needs as expressed by end-users such as Airbus, Areva/EDF and Thales.

Results on methodological enhancements are described

7.3. Coordination of heterogeneous Models of Computation as Domain-Specific Languages

Participants: Matias Vara Larsen, Julien Deantoni, Frédéric Mallet.

In the context of the collaborative ANR GEMOC project (9.2.1.2), we investigated the way the multiview approach generally promoted in Aoste could deal with analysis and simulation of systems specified using multiple heterogeneous languages. Coordinated use of heterogeneous domain specific languages (DSL) led to so-called globalization of modeling language. We wrote a chapter related to these concerns [50], as part of a book dedicated to the challenges of the field, gathering industrial and academic contributors.

This goal was achieved in two steps. First step consisted in specifying a language able to support appropriate information (*i.e.*, the one required for the coordination) in a *Language Behavioral Interface (LBI)*. Second step consisted in using the LBI to define coordination patterns from which the coordination of models can be automatically inferred. Design is supported by an heterogeneous simulation engine that has been developed and integrated in the Gemoc studio environment. Gemoc Studio, enhanced with our new research ideas, won the 9th execution tool contest at ...

We also developed MoCCML (Model of Concurrency and Communication Modeling Language), an imperative extension of the CCSL language in the form of constraint automata [28]. MoCCML defines the concurrent and communication part of the semantics of a language, and is used by the LBI to exhibit internal causalities and synchronizations. Finally, we defined a protocol combining the concurrency aspects and the execution functions (*i.e.*, the rewriting rules) so as to be able to develop, in a modular way, the whole behavioral semantics of a language [30], [31].

Our work this on coordination of heterogeneous languages produced two major results. The first one is the development of BCOoL (Behavioral Coordination Operator Language [33]). BCOoL is a language dedicated to the specification of coordination patterns between heterogeneous languages. It comes with a tool chain allowing the generation of the coordination given a BCOoL operator and specific models. Our second result is the development of an heterogeneous execution engine, integrated to Gemoc studio, to run conjointly different models [44]. Both works were mainly realized by Matias Vara Larsen, as part of his upcoming PhD.

7.4. SoC multiview (meta)modeling for performance, power, and thermal aspects

Participants: Amani Khecharem, Robert de Simone, Emilien Kofman, Julien Deantoni.

In the framework of the ANR HOPE project we progressed the definition of multiview metamodels for the design of Systems-on-Chip) (SoC systems integrating performance, power and thermal aspects. The main concern was to stress regularity and commonality between those views, each developed on "domains" defined as partitions of the original block diagram (clock domains, voltage domains, floorplans,...), and with finite state machine controllers setting the levels of these domains; links between distinct views are originally provided by laws of physics, but then usually identified on discrete allowable values by engineers. The application view, meant to provide typical use-cases to help dimension the SoC platform by abstract simulation, also fits in this framework. This methodological work was presented in the local forum SAME (Sophia-Antipolis MicroElectronics) [53]. It is supposed to work in two ways, both by allowing the application of analytic methods to compute an optimized mapping of application tasks onto platform resources, and then to translate these results towards sophisticated simulation environments (such as MCO Platform Architect by Synopsys or ACEplorer by Docea Power/Intel, both partners in the HOPE consortium) which consider non-functional aspects of power and thermal modeling in their simulation environments. The various approaches considered in Aoste to define mapping constraints and solve them algorithmically are presented elsewhere. All this should soon be reported in Amani Khacharem PhD document.

7.5. Networks-on-Board: between NoCs and rack connector buses

Participants: Amine Oueslati, Robert de Simone, Albert Savary, Emilien Kofman.

The recent paradigm of Massively Parallel Processor Arrays (MPPA), or more generally manycore Systems-on-Chip, rely on the existence of a high-throughput on-Chip Network (NoC) to interconnect the various cores and processing clusters. Despite its benefits, it requires that all components are put on the same die, and thus designed monolithically. On the other end, supercomputers are built by assembling racks or blades of processors, connected by fast buses (fast ethernet or infiniband usually), with low predictivity of throughput. A third, intermediate path is explored in the context of the FUI Clistine project, based on a notion of Network-on-Board (or Network-in-Package), aiming at the benefits of NoCs brought to the level of a single PCB board, where the various components can be assembled in a modular fashion. We consider the application of our previous expertise on modeling and analysis of NoC-based architecture, with their implications on the optimized mapping of dataflow models of applications onto such interconnects, to adapt them in this new context. The objective is to consider alternative network topologies, and to translate optimal mappings into the concrete network operations on a prototype implementation realized by SynergieCAD, the company heading the project. This topic reflects the PhD thesis of Amine Oueslati, and the engineering work of Albert Savary.

7.6. Solving AAA constraints analytically

Participants: Emilien Kofman, Dumitru Potop Butucaru, Thomas Carle, Raul Gorcitz, Robert de Simone, Mohamed Bergach, Amine Oueslati.

Given two abstract modeling descriptions, one of a dataflow process network for the application, one of a block diagram structure for the computing platform and its interconnects, together with cost functions for the elementary computations and communications, one is bound to seek optimal mappings pairing the two. Amongst all the possible techniques, an obvious one consists in solving constraint using general solvers (real, integer, or boolean constraint programming, SMT solvers, etc). Given the NP-hard nature of the problem, the issue here is to scale to the dimensions of realistic problems. We conducted extensive experiments on several case studies, with as extra objective the concern of studying how the formulation of constraints, or the exploitation of additional information (in concurrency or exclusion of tasks, structural symmetries,...) could impact favorably or negatively the process. Results were compiled in a publication [57].

In the framework of the PhD thesis of Mohamed Bergach, under CIFRE funding with Kontron Toulon, we studied how to adjust a radar application, that typically computes extensively FFT convolutions, on an hybrid CPU/GPU architecture such as IntelCore IvyBridge and Haswell processors. This approach works in two stages: first we considered how to implement a FFT redex as large as possible in exactly one core (either a CPU core or a GPU Execution Unit), so as to make full use of the local register memories and SIMD/vectorial instructions. Not by accident certainly FFT blocks of size exactly 8 and 16 respectively can so be fitted on a GPU (resp. CPU) block. This provides a new "compound" instruction, on which to build modularly and optimization the allocation of larger applications based on such basic block. This is fully described in Mohamed Bergach PhD document [16].

7.7. Stochastic extension of MARTE/CCSL for CPS modeling

Participant: Frédéric Mallet.

This work was conducted during the sabbatical period of Frédéric Mallet at ECNU Shanghai, in the context of the associated team FM4CPS (9.4.1.1).

As a declarative language, CCSL allows the specification of causal and temporal properties of systems expressed as constraints in a specific syntax. While each constraint reduces the set of possible behaviors, there may still be multiple (schedule) solutions, or none at all. When several solutions remain feasible, our TIMESQUARE tool allows to set up a resolving policy, to choose whether we want to attempt exploring exhaustively all these solutions, or else narrow the solution space according to an auxiliary criterion.

The extension of CCSL with stochastic features and probabilistic information is meant to help provide such an additional criterion, while modeling temporal constraints on the environment which are not necessarily well-known or controllable, specially in the domain of Cyber-Physical Systems. Then, such features should help reducing the set of possible behaviors, narrowing for instance to the most likely ones (in a formal quantitative meaning).

We are currently relying on UPPAAL SMC (Stochastic Model-Checking) toolset as prototype analyzer for the resulting specifications.

7.8. Coupling SystemC and FMI for co-simulation of Cyber-Physical Systems

Participants: Stefano Centomo, Julien Deantoni, Robert de Simone.

In the context of Stefano Centoma master internship, and in collaboration with his global supervisor Professor Davide Quaglia, from the University of Verona, we considered the possibility to build heterogenous, multi-physics co-simulation schemes for hybrid continuous-discrete Cyber-Physical systems. The first step consisted in extracting relevant interface information from IP component described in the SystemC language; it was naturally inspired from some of our former work. But currently IP-XACT is meant to address easy component assembly at the *structural* (static) level, and is not concerned with dynamical aspects of behavior simulation. This extension, and the proper combination with the FMI standard for its purpose, allowing hybrid and multiform co-simulation of SystemC components (and also others describing the continuous physical environment) are the next-step objective being currently tackled.

7.9. Code generation for time-triggered platforms based on Real-Time Scheduling

Participants: Dumitru Potop Butucaru, Raul Gorcitz, Yves Sorel.

We have continued this year the work on real-time scheduling and code generation for time-triggered platforms. Much of this work was carried out as part of a trilateral collaboration with Airbus DS and the CNES, which have funded an (onerous) TTEthernet-based test platform and partly funded the post-doctorate of Raul Gorcitz. The remainder of Raul Gorcitz' post-doc has been funded by the ITEA3 Assume project.

This year, the objective has been to allow code generation on an industry-grade platform comprising ARINC 653-based computers connected through a TTEthernet network. The novelty with respect to previous years comes from the time-triggered TTEthernet network, whose scheduling properties raise new problems. Unlike in classical field buses, resource reservation in a TTEthernet network is done at the level of directed links (physical wires that connect routers and end stations). Each of these links is controlled by an arbiter that determines the scheduling of both time-triggered data transfers and control messages needed to ensure the global time synchronization. This year we have built a model of the TTEthernet network allowing precise real-time scheduling, and worked on code generation aspects. We expect to have a fully running prototype in the next 2 months, and to demonstrate it to our funders. Relevant publications are [18], [38].

For teaching purposes and to achieve a finer understanding of ARINC 653-based operating systems, we have also developed an implementation of the standard on inexpensive RaspberryPi platforms, and published a scientific vulgarization paper [55].

7.10. Real-time systems compilation

Participants: Dumitru Potop Butucaru, Keryan Didier, Mihail Asavoae.

This research line develops over various results of the team over the years, its aim being to develop fully automatic implementation flows going fully automatically from functional and non-functional specification to correct and efficient running implementation. We advocate for a real-time systems compilation approach that combines aspects of both real-time scheduling and compilation of both classical and synchronous languages. Like a classical compiler such as GCC, a real-time systems compiler should use fast and efficient scheduling and code generation heuristics, to ensure scalability. Similarly, it should provide traceability support under the form of informative error messages enabling an incremental trial-and-error design style, much like that of classical application software. This is more difficult than in a classical compiler, given the complexity of the transformation flow (creation of tasks, allocation, scheduling, synthesis of communication and synchronization code, etc.), and requires a full formal integration along the whole flow, including the crucial issue of correct hardware abstraction. A real-time systems compiler should perform precise, conservative timing

accounting along the whole scheduling and code generation flow, allowing it to produce safe and tight real-time guarantees. More generally, and unlike in classical compilers, the allocation and scheduling algorithms must take into account a variety of non-functional requirements, such as real-time constraints, criticality, partitioning, preemptability, allocation constraints, etc. As the accent is put on the respect of requirements (as opposed to optimization of a metric, like in classical compilation), resulting scheduling problems are quite different.

We are currently building such a real-time systems compiler, called Lopht. The construction of the Lopht tool, which takes into account complex functional and non-functional specifications is discussed in the corresponding section and in [17].

This year, we have initiated work on two fundamental topics. The first one is sound architecture abstraction – ensuring that the platform models used for real-time scheduling and code generation are conservative abstractions of the real hardware and basic software, allowing the generation of implementations that are functionally and non-functionally correct. This work is performed in the framework of the LEOC Capacites project, which funds the post-doc of Mihail Asavae. The second line of work aims at formally proving that the output of Lopht is correct with respect to its input models (including functional specification and platform model). This work is performed in the ITEA3 Assume project, which funds the PhD thesis of Keryan Didier. Together with the Parkas team-project we have also considered the implementation of mixed-criticality systems [26].

7.11. Uniprocessor Real-Time Scheduling

Participants: Mamadou Diallo, Yves Sorel, Walid Talaboulma, Robert Davis.

In the context of the master internship of Mamadou Diallo we implemented the offline time trigger scheduler proposed in his PhD thesis by Falou Ndoeye on a development board based on an ARM Cortex M4. We used this ARM version since it is better suited to embedded systems, since more predictable, than the ARM 7 we used last year. Especially, it allows to determine more accurately the cost of the scheduler and of the preemptions we use in our offline schedulability analysis. We remind that the schedulability analysis provides a scheduling table which is exploited by the scheduler during the real-time execution of the tasks. This approach allows a low and fixed cost for the scheduler and the preemptions whereas these costs are variable in the case of classical online schedulers. For several task sets we compared the timing diagrams predicted by the schedulability analysis with the real-time timing diagrams measured on the ARM Cortex M4. It turns out that those timings are very close, as expected.

A new direction opened with the arrival of Rob Davis was to consider by studying the impact of the non-preemptivity constraints on the optimality of the schedulers [37], or by considering fixed priorities while scheduling messages in the context of Control Area Networks [36].

7.12. Multiprocessor Real-Time Scheduling

Participants: Aderraouf Benyahia, Laurent George, Salah Eddine Saidi, Yves Sorel, Robert Davis, Liliana Cucu.

In the context of the PhD thesis of Salah Eddine Saidi we considered the co-simulation of several process models specified in continuous time and several controllers models specified in discrete time according to a real-time hardware in the loop approach. These models specified with different tools such as Simulink, AMESim, Modelica, etc., cooperate according to the FMI standard. They are translated in a dataflow graph that is compliant with the conditioned repetitive dataflow model of our AAA methodology for functional specification. Each model considers the feed-through function as well as the functions which depend of the state, and the state computation itself. In order to meet the real-time constraints of such complex co-simulation we need to execute them on multicore platforms. We studied the limitations of greedy and local search distributed real-time scheduling heuristics we developed in the past for control applications. The first limitation is related to the FMI standard which requires that the functions belonging to a model are allocated to the same core. We first try to introduce additional semaphores in the real-time code generated automatically

to avoid these situations. Unfortunately, this solution decreases significantly the acceleration brought by the multicore. Therefore, we started to investigate graph based techniques that add non directed edges to specify the FMI relation and search solutions where some non oriented edges can be oriented to minimize locally the makespan.

In the context of the master internship of Mamadou Diallo we studied the possibilities to extend the offline time trigger scheduler implemented on a uniprocessor to the multiprocessor case. Since the embedded board based on the ARM Cortex M4 we utilize features an ethernet interface, we conducted several experimentations on ethernet switches to measure the end-to-end communication time between several real-time tasks running on such boards with such schedulers.

We completed the work on the gateway with modeling languages for certified code generation carried out in the P FUI project 9.2.2 which ended in June 2015. Mainly, we tested the P modeling language to SynDEx gateway on four industrial use cases provided by AdaCore, Continental and Aboard Engineering. We specified these applications with the P language and translated them in the SynDEx format. With SynDEx we analysed the schedulability and automatically generated the corresponding code for an Intel 8 cores Xeon ES-1620v2 3.70Ghz. For these applications ranging from 103 to 1403 bloks we obtained an acceleration factor equal to the number of cores.

Thanks in part to the arrival of Rob Davis, our team has participated to the proposition of a new framework in the context of multicore platforms: *Multicore Response Time Analysis framework* [34]. This proposal was made in close collaboration with academic partners such as the University of Luxembourg, Verimag and ISEP Porto. The framework is extensible to different multicore architectures, with various types and arrangements of local memory, and different arbitration policies for the common interconnects. The MRTA framework provides a general approach to timing verification for multicore systems, parametric in the hardware configuration, and so can be used architectural design stage to compare the guaranteed levels of performance that can be obtained with different hardware configurations. The MRTA framework decouples response time analysis from a reliance on context independent WCET values. Instead, the analysis specifies response times directly according to requirements on different hardware resources.

7.13. Probabilistic and statistical temporal analysis

Participants: Liliana Cucu, Robert Davis, Adriana Gogonel, Walid Talaboulma, Dorin Maxim, Cristian Maxim.

Real-time constraint guarantees require worst-case reasoning to provide sound solutions. We have proposed to define and use worst-case reasoning in different contexts: optimal scheduling algorithms, response time analysis, estimation of worst-case execution times. These results have laid the foundations for certifiable probabilistic solutions to real-time systems.

In particular, we have studied the probabilistic response time analysis for systems with multiple probabilistic parameters, either by using bounds based on real-time calculus, extreme value theory, direct calculation or in a context of component-based systems. Generally, probabilistic methods have high complexity cost; using upper-bounds for the input probability distributions we provide conservative(safe) results faster. Worst-case reasoning is also provided for the statical estimation of a task probabilistic worst-case execution time.

Results were published in [22], [24], [58], [56], [42], [46], [23], [42], [43], [40]

7.14. Parametric and Non-Parametric Statistics for Program Performance Analysis and Comparison

Participant: Sid Touati.

This research activity is a continuation of our joint research effort with Julien Worms, Assistant Professor at University of Versailles Saint-Quentin (UVSQ), dealing with statistical program performance analysis and comparison, in presence of performance variability. In the previous study (called Speedup-Test), we gave a rigorous statistical methodology for analysis of program speedups based on mean or median performance metrics: execution time, energy consumption, etc. However mean or median observed performances do not always reflect the user's feeling of performance, especially when the performances are really unstable. In the current study, we propose additional precise performance metrics, based on performance modeling using gaussian mixtures. We explore the difference between parametric and non parametric statistics applied on program performance analysis. Our additional statistical metrics for analysing and comparing program performances give to the user more precise decision tools to select best code versions, not necessarily based on mean or median numbers. Also, we provide a new metric to estimate performance variability based on gaussian mixture model. Our statistical methods are implemented in R, and distributed as open source code. A research report is under completion, before submission as article.

ARIC Project-Team

7. New Results

7.1. Floating-point Arithmetic

7.1.1. *On the maximum relative error when computing integer powers by iterated multiplications in floating-point arithmetic*

We improve the usual relative error bound for the computation of x^n through iterated multiplications by x in binary floating-point arithmetic. The obtained error bound is only slightly better than the usual one, but it is simpler. We also discuss the more general problem of computing the product of n terms. [5]

7.1.2. *Formally verified certificate checkers for hardest-to-round computation*

In order to derive efficient and robust floating-point implementations of a given function f , it is crucial to compute its hardest-to-round points, i.e. the floating-point numbers x such that $f(x)$ is closest to the midpoint of two consecutive floating-point numbers. Depending on the floating-point format one is aiming at, this can be highly computationally intensive. In this paper, we show how certificates based on Hensel's lemma can be added to an algorithm using lattice basis reduction so that the result of a computation can be formally checked in the Coq proof assistant. [7]

7.1.3. *On the error of Computing $ab + cd$ using Cornea, Harrison and Tang's method*

In their book, Scientific Computing on the Itanium, Cornea et al. (2002) introduce an accurate algorithm for evaluating expressions of the form $ab + cd$ in binary floating-point arithmetic, assuming an FMA instruction is available. They show that if p is the precision of the floating-point format and if $u = 2^{-p}$, the relative error of the result is of order u . We improve their proof to show that the relative error is bounded by $2u + 7u^2 + 6u^3$. Furthermore, by building an example for which the relative error is asymptotically (as $p \rightarrow \infty$ or, equivalently, as $u \rightarrow 0$) equivalent to $2u$, we show that our error bound is asymptotically optimal. [8]

7.1.4. *Improved error bounds for floating-point products and Horner's scheme*

Let u denote the relative rounding error of some floating-point format. Recently it has been shown that for a number of standard Wilkinson-type bounds the typical factors $\gamma_k := ku/(1-ku)$ can be improved into ku , and that the bounds are valid without restriction on k . Problems include summation, dot products and thus matrix multiplication, residual bounds for LU- and Cholesky-decomposition, and triangular system solving by substitution. In this note we show a similar result for the product $\prod_{i=0}^k x_i$ of real and/or floating-point numbers x_i , for computation in any order, and for any base $\beta \geq 2$. The derived error bounds are valid under a mandatory restriction of k . Moreover, we prove a similar bound for Horner's polynomial evaluation scheme. [9]

7.1.5. *Comparison between binary and decimal floating-point numbers*

In collaboration with Christoph Lauter and Marc Mezzarobba (LIP6 laboratory, Paris), Nicolas Brisebarre and Jean-Michel Muller introduce an algorithm to compare a binary floating-point (FP) number and a decimal FP number, assuming the "binary encoding" of the decimal formats is used, and with a special emphasis on the basic interchange formats specified by the IEEE 754-2008 standard for FP arithmetic. It is a two-step algorithm: a first pass, based on the exponents only, quickly eliminates most cases, then, when the first pass does not suffice, a more accurate second pass is performed. They provide an implementation of several variants of our algorithm, and compare them [26].

7.2. Lattices: algorithms and cryptology

7.2.1. *Linearly Homomorphic Encryption from DDH*

We design a linearly homomorphic encryption scheme whose security relies on the hardness of the decisional Diffie-Hellman problem. Our approach requires some special features of the underlying group. In particular, its order is unknown and it contains a subgroup in which the discrete logarithm problem is tractable. Therefore, our instantiation holds in the class group of a non maximal order of an imaginary quadratic field. Its algebraic structure makes it possible to obtain such a linearly homomorphic scheme whose message space is the whole set of integers modulo a prime p and which supports an unbounded number of additions modulo p from the ciphertexts. A notable difference with previous works is that, for the first time, the security does not depend on the hardness of the factorization of integers. As a consequence, under some conditions, the prime p can be scaled to fit the application needs. [13]

7.2.2. *Secure Efficient History-Hiding Append-Only Signatures in the Standard Model*

As formalized by Kiltz et al. (ICALP'05), append-only signatures (AOS) are digital signature schemes where anyone can publicly append extra message blocks to an already signed sequence of messages. This property is useful, e.g., in secure routing, in collecting response lists, reputation lists, or petitions. Bethencourt, Boneh and Waters (NDSS'07) suggested an interesting variant, called history-hiding append-only signatures (HH-AOS), which handles messages as sets rather than ordered tuples. This HH-AOS primitive is useful when the exact order of signing needs to be hidden. When free of subliminal channels (i.e., channels that can tag elements in an undetectable fashion), it also finds applications in the storage of ballots on an electronic voting terminals or in other archival applications (such as the record of petitions, where we want to hide the influence among messages). However, the only subliminal-free HH-AOS to date only provides heuristic arguments in terms of security: Only a proof in the idealized (non-realizable) random oracle model is given. This paper provides the first HH-AOS construction secure in the standard model. Like the system of Bethencourt et al., our HH-AOS features constant-size public keys, no matter how long messages to be signed are, which is atypical (we note that secure constructions often suffer from a space penalty when compared to their random-oracle-based counterpart). As a second result, we show that, even if we use it to sign ordered vectors as in an ordinary AOS (which is always possible with HH-AOS), our system provides considerable advantages over existing realizations. As a third result, we show that HH-AOS schemes provide improved identity-based ring signatures (i.e., in prime order groups and with a better efficiency than the state-of-the-art schemes). [17]

7.2.3. *Compactly Hiding Linear Spans: Tightly Secure Constant-Size Simulation-Sound QA-NIZK Proofs and Applications*

Quasi-adaptive non-interactive zero-knowledge (QA-NIZK) proofs is a powerful paradigm, suggested recently by Jutla and Roy (Asiacrypt'13), which is motivated by the Groth-Sahai seminal techniques for efficient non-interactive zero-knowledge (NIZK) proofs. In this paradigm, the common reference string may depend on specific language parameters, a fact that allows much shorter proofs in important cases. It even makes certain standard model applications competitive with the Fiat-Shamir heuristic in the Random Oracle idealization (such QA-NIZK proofs were recently optimized to constant size by Jutla and Roy (Crypto'14) and Libert et al. (Eurocrypt'14) for the important case of proving that a vector of group elements belongs to a linear subspace). While, e.g., the QA-NIZK arguments of Libert et al. provide unbounded simulation-soundness and constant proof length, their simulation-soundness is only loosely related to the underlying assumption (with a gap proportional to the number of adversarial queries) and it is unknown how to alleviate this limitation without sacrificing efficiency. Here, we deal with the basic question of whether and to what extent we can simultaneously optimize the proof size and the tightness of security reductions, allowing for important applications with tight security (which are typically to date quite lengthy) to be of shorter size. In this paper, we resolve this question by describing a novel simulation-sound QA-NIZK argument showing that a vector $v \in G^n$ belongs to a subspace of rank $t < n$ using a constant number of group elements. Unlike previous constant-size QA-NIZK proofs of such statements, the unbounded simulation-soundness of our system is nearly tightly related (i.e., the reduction only loses a factor proportional to the security parameter) to the standard Decision

Linear assumption. To show simulation-soundness in the constrained context of tight reductions, we employ a number of techniques, and explicitly point at a technique – which may be of independent interest – of hiding the linear span of a structure-preserving homomorphic signature (which is part of an OR proof). As an application, we design a public-key cryptosystem with almost tight CCA2-security in the multi-challenge, multiuser setting with improved length (asymptotically optimal for long messages). We also adapt our scheme to provide CCA security in the key-dependent message scenario (KDM-CCA2) with ciphertext length reduced by 75% when compared to the best known tightly secure KDM-CCA2 system so far. [18]

7.2.4. Short Group Signatures via Structure-Preserving Signatures: Standard Model Security from Simple Assumptions

Group signatures are a central cryptographic primitive which allows users to sign messages while hiding their identity within a crowd of group members. In the standard model (without the random oracle idealization), the most efficient constructions rely on the Groth-Sahai proof systems (Eurocrypt’08). The structure-preserving signatures of Abe et al. (Asiacrypt’12) make it possible to design group signatures based on well-established, constant-size number theoretic assumptions (a.k.a. “simple assumptions”) like the Symmetric eXternal Diffie-Hellman or Decision Linear assumptions. While much more efficient than group signatures built on general assumptions, these constructions incur a significant overhead w.r.t. constructions secure in the idealized random oracle model. Indeed, the best known solution based on simple assumptions requires 2.8 kB per signature for currently recommended parameters. Reducing this size and presenting techniques for shorter signatures are thus natural questions. In this paper, our first contribution is to significantly reduce this overhead. Namely, we obtain the first fully anonymous group signatures based on simple assumptions with signatures shorter than 2 kB at the 128-bit security level. In dynamic (resp. static) groups, our signature length drops to 1.8 kB (resp. 1 kB). This improvement is enabled by two technical tools. As a result of independent interest, we first construct a new structure-preserving signature based on simple assumptions which shortens the best previous scheme by 25%. Our second tool is a new method for attaining anonymity in the strongest sense using a new CCA2-secure encryption scheme which is simultaneously a Groth-Sahai commitment. [19]

7.2.5. Implementing Candidate Graded Encoding Schemes from Ideal Lattices

Multilinear maps have become popular tools for designing cryptographic schemes since a first approximate realisation candidate was proposed by Garg, Gentry and Halevi (GGH). This construction was later improved by Langlois, Stehlé and Steinfeld who proposed GGHLite which offers smaller parameter sizes. In this work, we provide the first implementation of such approximate multilinear maps based on ideal lattices. Implementing GGH-like schemes naively would not allow instantiating it for non-trivial parameter sizes. We hence propose a strategy which reduces parameter sizes further and several technical improvements to allow for an efficient implementation. In particular, since finding a prime ideal when generating instances is an expensive operation, we show how we can drop this requirement. We also propose algorithms and implementations for sampling from discrete Gaussians, for inverting in some Cyclotomic number fields and for computing norms of ideals in some Cyclotomic number rings. Due to our improvements we were able to compute a multilinear jigsaw puzzle for $\kappa = 52$ (resp. $\kappa = 38$) and $\lambda = 52$ (resp. $\lambda = 80$). [10]

7.2.6. Improved security proofs in lattice-based cryptography: using the Rényi divergence rather than the statistical distance

The Rényi divergence is a mean to measure the closeness of two distributions. We show that it can often be used as an alternative to the statistical distance in security proofs for lattice-based cryptography. Using the Rényi divergence is particularly suited for security proofs of primitives in which the attacker is required to solve a search problem (e.g., forging a signature). We show that it may also be used in the case of distinguishing problems (e.g., semantic security of encryption schemes), when they enjoy a public sampleability property. The techniques lead to security proofs for schemes with smaller parameters. [11]

7.2.7. Fully Secure Functional Encryption for Inner Products, from Standard Assumptions

Functional encryption is a modern public-key paradigm where a master secret key can be used to derive sub-keys SKF associated with certain functions F in such a way that the decryption operation reveals $F(M)$, if

M is the encrypted message, and nothing else. Recently, Abdalla et al. gave simple and efficient realizations of the primitive for the computation of linear functions on encrypted data: given an encryption of a vector y over some specific base ring, a secret key SK_x for the vector x allows computing $\langle x, y \rangle$. Their technique surprisingly allows for instantiations under standard assumptions, like the hardness of the Decision Diffie-Hellman (DDH) and Learning-with-Errors (LWE) problems. Their constructions, however, are only proved secure against selective adversaries, which have to declare the challenge messages M_0 and M_1 at the outset of the game. In this paper, we provide constructions that provably achieve security against more realistic adaptive attacks (where the messages M_0 and M_1 may be chosen in the challenge phase, based on the previously collected information) for the same inner product functionality. Our constructions are obtained from hash proof systems endowed with homomorphic properties over the key space. They are (almost) as efficient as those of Abdalla et al. and rely on the same hardness assumptions. In addition, we obtain a solution based on Paillier's composite residuosity assumption, which was an open problem even in the case of selective adversaries. We also propose LWE-based schemes that allow evaluation of inner products modulo a prime p , as opposed to the schemes of Abdalla et al. that are restricted to evaluations of integer inner products of short integer vectors. We finally propose a solution based on Paillier's composite residuosity assumption that enables evaluation of inner products modulo an RSA integer $N = pq$. We demonstrate that the functionality of inner products over a prime field is very powerful and can be used to construct bounded collusion FE for all circuits. [23]

7.2.8. Fully Homomorphic Encryption over the Integers Revisited

Two main computational problems serve as security foundations of current fully homomorphic encryption schemes: Regev's Learning With Errors problem (LWE) and Howgrave-Graham's Approximate Greatest Common Divisor problem (AGCD). Our first contribution is a reduction from LWE to AGCD. As a second contribution, we describe a new AGCD-based fully homomorphic encryption scheme, which outperforms all prior AGCD-based proposals: its security does not rely on the presumed hardness of the so-called Sparse Subset Sum problem, and the bit-length of a ciphertext is only $\tilde{O}\lambda$, where λ refers to the security parameter. [15]

7.2.9. Cryptanalysis of the Multilinear Map over the Integers

We describe a polynomial-time cryptanalysis of the (approximate) multilinear map of Coron, Lepoint and Tibouchi (CLT). The attack relies on an adaptation of the so-called zeroizing attack against the Garg, Gentry and Halevi (GGH) candidate multilinear map. Zeroizing is much more devastating for CLT than for GGH. In the case of GGH, it allows to break generalizations of the Decision Linear and Subgroup Membership problems from pairing-based cryptography. For CLT, this leads to a total break: all quantities meant to be kept secret can be efficiently and publicly recovered. [14]

7.2.10. Cryptanalysis of Gu's ideal multilinear map

In March, 2015 Gu Chunsheng proposed a candidate ideal multilinear map [eprint 2015/269]. An ideal multilinear map allows to perform as many multiplications as desired, while in k -multilinear maps like GGH [EC 2013] or CLT [CR2013, CR2015] one can perform at most a predetermined number k of multiplications. In this note, we show that the extraction Multilinear Computational Diffie-Hellman problem (ext-MCDH) associated to Gu's map can be solved in polynomial-time: this candidate ideal multilinear map is insecure. We also give intuition on why we think that the two other ideal multilinear maps proposed by Gu in [eprint 2015/269] are not secure either. [39]

7.2.11. Worst-case to average-case reductions for module lattices

Most lattice-based cryptographic schemes are built upon the assumed hardness of the Short Integer Solution (SIS) and Learning With Errors (LWE) problems. Their efficiencies can be drastically improved by switching the hardness assumptions to the more compact Ring-SIS and Ring-LWE problems. However, this change of hardness assumptions comes along with a possible security weakening: SIS and LWE are known to be at least as hard as standard (worst-case) problems on euclidean lattices, whereas Ring-SIS and Ring-LWE are only known to be as hard as their restrictions to special classes of ideal lattices, corresponding to ideals of some

polynomial rings. In this work, we define the Module-SIS and Module-LWE problems, which bridge SIS with Ring-SIS, and LWE with Ring-LWE, respectively. We prove that these average-case problems are at least as hard as standard lattice problems restricted to module lattices (which themselves generalize arbitrary and ideal lattices). As these new problems enlarge the toolbox of the lattice-based cryptographer, they could prove useful for designing new schemes. Importantly, the worst-case to average-case reductions for the module problems are (qualitatively) sharp, in the sense that there exist converse reductions. This property is not known to hold in the context of Ring-SIS/Ring-LWE: Ideal lattice problems could reveal easy without impacting the hardness of Ring-SIS/Ring-LWE. [6]

7.2.12. Reducing Communication Overhead of the Subset Difference Scheme

In Broadcast Encryption (BE) systems like Pay-TV, AACS, online content sharing and broadcasting, reducing the header length (communication overhead per session) is of practical interest. The Subset Difference (SD) scheme due to Naor-Naor-Lotspiech (NNL) is the most popularly used BE scheme. This work introduced the (a, b, γ) augmented binary tree subset difference $((a, b, \gamma)$ -ABTSD) scheme which is a generalization of the NNL-SD scheme. By varying the parameters (a, b, γ) , it is possible to obtain $O(n \log n)$ different schemes. In addition to the underlying binary tree structure of the NNL-SD scheme, the new scheme uses an additional binary tree structure of height a augmented with each internal node. The SD subsets in this scheme arise due to nodes that are at a distance at most b from each other. In the augmented tree of height a , at most c leaves are considered together in creating the SD subsets for the scheme. The average header length achieved by the new schemes is smaller than all known schemes having the same decryption time as that of the NNL-SD scheme and achieving non-trivial trade-offs between the user storage and the header size. The amount of key material that a user is required to store increases. For the earlier mentioned applications, reducing header size and achieving fast decryption is perhaps more of a concern than the user storage

7.3. Algebraic computing and high performance kernels

7.3.1. Complexity of the F5 Gröbner basis algorithm

We study the complexity of Gröbner bases computation, in particular in the generic situation where the variables are in simultaneous Noether position with respect to the system. We give a bound on the number of polynomials of degree d in a Gröbner basis computed by Faugère's F5 algorithm (2002) in this generic case for the grevlex ordering (which is also a bound on the number of polynomials for a reduced Gröbner basis, independently of the algorithm used). Next, we analyse more precisely the structure of the polynomials in the Gröbner bases with signatures that F5 computes and use it to bound the complexity of the algorithm. Our estimates show that the version of F5 we analyse, which uses only standard Gaussian elimination techniques, outperforms row reduction of the Macaulay matrix with the best known algorithms for moderate degrees, and even for degrees up to the thousands if Strassen's multiplication is used. The degree being fixed, the factor of improvement grows exponentially with the number of variables. [1]

7.3.2. Faster Algorithms for Multivariate Interpolation with Multiplicities and Simultaneous Polynomial Approximations

The interpolation step in the Guruswami-Sudan algorithm is a bivariate interpolation problem with multiplicities commonly solved in the literature using either structured linear algebra or basis reduction of polynomial lattices. This problem has been extended to three or more variables; for this generalization, all fast algorithms proposed so far rely on the lattice approach. In this work, we reduce this multivariate interpolation problem to a problem of simultaneous polynomial approximations, which we solve using fast structured linear algebra. This improves the best known complexity bounds for the interpolation step of the list-decoding of Reed-Solomon codes, Parvaresh-Vardy codes, and folded Reed-Solomon codes. In particular, for Reed-Solomon list-decoding with re-encoding, our approach has complexity $\tilde{O}(\ell^{\omega-1} m^2 (n-k))$, where ℓ, m, n, k are the list size, the multiplicity, the number of sample points and the dimension of the code, and ω is the exponent of linear algebra; this accelerates the previously fastest known algorithm by a factor of ℓ/m . [3]

7.3.3. Recursion based parallelization of exact dense linear algebra routines for Gaussian elimination

We present block algorithms and their implementation for the parallelization of sub-cubic Gaussian elimination on shared memory architectures. Contrarily to the classical cubic algorithms in parallel numerical linear algebra, we focus here on recursive algorithms and coarse grain parallelization. Indeed, sub-cubic matrix arithmetic can only be achieved through recursive algorithms making coarse grain block algorithms perform more efficiently than fine grain ones. This work is motivated by the design and implementation of dense linear algebra over a finite field, where fast matrix multiplication is used extensively and where costly modular reductions also advocate for coarse grain block decomposition. We incrementally build efficient kernels, for matrix multiplication first, then triangular system solving, on top of which a recursive PLUQ decomposition algorithm is built. We study the parallelization of these kernels using several algorithmic variants: either iterative or recursive and using different splitting strategies. Experiments show that recursive adaptive methods for matrix multiplication, hybrid recursive-iterative methods for triangular system solve and tile recursive versions of the PLUQ decomposition, together with various data mapping policies, provide the best performance on a 32 cores NUMA architecture. Overall, we show that the overhead of modular reductions is more than compensated by the fast linear algebra algorithms and that exact dense linear algebra matches the performance of full rank reference numerical software even in the presence of rank deficiencies. [4]

7.3.4. Computing the Rank Profile Matrix

The row (resp. column) rank profile of a matrix describes the staircase shape of its row (resp. column) echelon form. In an ISSAC'13 paper, we proposed a recursive Gaussian elimination that can compute simultaneously the row and column rank profiles of a matrix as well as those of all of its leading sub-matrices, in the same time as state of the art Gaussian elimination algorithms. Here we first study the conditions making a Gaussian elimination algorithm reveal this information. Therefore, we propose the definition of a new matrix invariant, the rank profile matrix, summarizing all information on the row and column rank profiles of all the leading sub-matrices. We also explore the conditions for a Gaussian elimination algorithm to compute all or part of this invariant, through the corresponding PLUQ decomposition. As a consequence, we show that the classical iterative CUP decomposition algorithm can actually be adapted to compute the rank profile matrix. Used, in a Crout variant, as a base-case to our ISSAC'13 implementation, it delivers a significant improvement in efficiency. Second, the row (resp. column) echelon form of a matrix are usually computed via different dedicated triangular decompositions. We show here that, from some PLUQ decompositions, it is possible to recover the row and column echelon forms of a matrix and of any of its leading sub-matrices thanks to an elementary post-processing algorithm. [16]

7.3.5. Formulas for Continued Fractions. An Automated Guess and Prove Approach

We describe a simple method that produces automatically closed forms for the coefficients of continued fractions expansions of a large number of special functions. The function is specified by a non-linear differential equation and initial conditions. This is used to generate the first few coefficients and from there a conjectured formula. This formula is then proved automatically thanks to a linear recurrence satisfied by some remainder terms. Extensive experiments show that this simple approach and its straightforward generalization to difference and q -difference equations capture a large part of the formulas in the literature on continued fractions. [20]

7.3.6. Algebraic Diagonals and Walks

The diagonal of a multivariate power series F is the univariate power series $\text{Diag } F$ generated by the diagonal terms of F . Diagonals form an important class of power series; they occur frequently in number theory, theoretical physics and enumerative combinatorics. We study algorithmic questions related to diagonals in the case where F is the Taylor expansion of a bivariate rational function. It is classical that in this case $\text{Diag } F$ is an algebraic function. We propose an algorithm that computes an annihilating polynomial for $\text{Diag } F$. Generically, it is its minimal polynomial and is obtained in time quasi-linear in its size. We show that this minimal polynomial has an exponential size with respect to the degree of the input rational function. We then

address the related problem of enumerating directed lattice walks. The insight given by our study leads to a new method for expanding the generating power series of bridges, excursions and meanders. We show that their first N terms can be computed in quasi-linear complexity in N , without first computing a very large polynomial equation. [12]

ATEAMS Project-Team

6. New Results

6.1. Faster Immutable Data Structures for the JVM

Immutable data structures involve copying when updating. Efficient implementations use persistent data-structures, so that most of the unchanged data is shared between the copies. Existing libraries for such data structures in the context of the Java virtual machine (JVM), such as the data structures in Clojure and Scala, are based on Hash Array-Mapped Tries (HAMTs), which provide efficient insertion and concatenation operations for persistent maps and sets. In [37] Steindorfer and Vinju presented additional optimisation which allow such operations to be up to 28 times faster than in the Clojure and Scala libraries. Furthermore, the cost of equality checking of such data structures is lower as well. All this, without incurring additional memory.

6.2. Automated Measurement and Analysis of Open Source Software

Deciding whether an open source software (OSS) meets the required standards for adoption in terms of quality, maturity, activity of development and user support is not a straightforward process. It involves analysing various sources of information, including the project's source code repositories, communication channels, and bug tracking systems. OSSMETER extends state-of-the-art techniques in the field of automated analysis and measurement of open-source software (OSS), and develops a platform that supports decision makers in the process of discovering, comparing, assessing and monitoring the health, quality, impact and activity of opensource software. To achieve this, OSSMETER computes trustworthy quality indicators by performing advanced analysis and integration of information from diverse sources including the project metadata, source code repositories, communication channels and bug tracking systems of OSS projects [29], [26]

This result comes from intensive collaboration in the FP7 STREP project "OSSMETER". The ATEAMS contribution is focused around source code metrics and activity analysis for Java and PHP.

6.3. Modular Interpreters for the Masses

Object Algebras [46] are new design pattern for increased modularity and extensibility of tree based, abstract data types. By modelling the abstract syntax of a language as a generic factory interface, implementations of this interface provide multiple semantics of the data. For instance, one can define evaluation, type checking and pretty printing of the abstract syntax fully modularly. Additionally, the pattern allows syntax extension: adding a new constructor to the datatype, and modularly extending any existing interpretations to deal with the construct. The same interpretation of different constructs, however, might involve different kinds of context information. For instance, evaluation of arithmetic expressions does not require any context information, but evaluation of variables and binders requires an environment. In [34], Inostroza and Van der Storm introduce a simple, modular, and type safe technique to allow such interpretations to be composed anyway. It is based on lifting one interpreter to implicitly propagate the context information it does not require, so that the signatures of the interpreters become compatible. As a result, semantic definitions of language modules do not have to anticipate all kinds of context information that might be required by other modules with which it might be composed. The technique is simple, does not sacrifice separate compilation, is easy to automate, and works in mainstream languages. It provides a first step towards a foundation for defining language by assembling modular building blocks.

6.4. One Parser to Rule Them All

Parsing realistic languages requires much more than just a parsing algorithm. Different kinds of language require advanced disambiguation, operator priorities, off-side rule checking, whitespace dependence or data dependence. In [25], Afroozeh and Izmaylova showed how most of these concerns are actually instances of data dependent parsing: the parsing process depends on the value of previously parsed input. They provided an encoding of indentation sensitive parsing, operator precedence and parsing in the presence of preprocessor directives, to a simple, data dependent core language which is executed using the general parsing algorithm GLL. By exposing the data dependent machinery at the level of the grammar formalism, this opens up a range of possibilities for custom parsing aspects, and provides a clear semantics for existing concerns like disambiguation.

6.5. A Pattern-Based Game Mechanics Design Assistant

Video game designers iteratively improve player experience by play testing game software and adjusting its design. Deciding how to improve gameplay, however, is difficult and time-consuming: designers lack an effective means for exploring decision alternatives and modifying a game's mechanics. In [35], Van Rozen presented the Mechanics Pattern Language (MPL) for encoding common game economy structures and design intent, and a Mechanics Design Assistant (MeDeA) for analyzing, explaining, understanding existing mechanics, and generating, filtering, exploring and applying design alternatives for modifying mechanics. As a result, game designers' productivity and game quality is increased by providing feedback and design alternatives early in the development cycle. Furthermore, the game economy modifications are applied at runtime using the MicroMachinations library, so that the effect of changes can be immediately experienced.

CAIRN Project-Team

7. New Results

7.1. Reconfigurable Architecture Design

7.1.1. Design Flow and Run-Time Management for Compressed FPGA Configurations

Participants: Olivier Sentieys, Christophe Huriaux.

Almost since the creation of the first SRAM-based FPGAs there has been a desire to explore the benefits of partially reconfiguring a portion of an FPGA at run-time while the remainder of design functionality continues to operate uninterrupted. Currently, the use of partial reconfiguration imposes significant limitations on the FPGA design: reconfiguration regions must be constrained to certain shapes and sizes and, in many cases, bitstreams must be precompiled before application execution depending on the precise region of the placement in the fabric. We developed an FPGA architecture that allows for seamless translation of partially-reconfigurable regions, even if the relative placement of fixed-function blocks within the region is changed.

In [42] we proposed a design flow for generating compressed configuration bit-streams abstracted from their final position on the logic fabric. Those configurations can then be decoded and finalized in real-time and at run-time by a dedicated reconfiguration controller to be placed at a given physical location. The VTR framework has been expanded to include bit-stream generation features. A bit-stream format is proposed to take part of our approach and the associated decoding architecture was designed. We analyzed the compression induced by our coding method and proved that compression ratios of at least $2.5\times$ can be achieved on the 20 largest MCNC benchmarks. The introduction of clustering which aggregates multiple routing resources together showed compression ratio up to a factor of $10\times$, at the cost of a more complex decoding step at runtime. The VBS approach can provide increased online relocation capabilities using a decoding algorithm capable of decoding the VBS on-the-fly during the task migration.

7.1.2. Run-Time Approximation under Performance Constraints in OFDM Wireless Receivers

Participants: Olivier Sentieys, Fernando Cladera.

Mobile wireless channels are characterized by time-varying multipath propagation, noise, and interference effects. To cope with these rapid variations of channel parameters, wireless receivers are designed with a significant performance margin to be able to reach a given link quality (BER - Bit Error Rate), even for the worst-case channel conditions. Indeed, one of the steps during the design phase is the choice of the architecture bit-width, and the smallest wordlength that ensures the correct behaviour of the receiver is usually chosen. In [39], an adaptive precision OFDM receiver is proposed. Significant energy savings come from varying at run time processing bit-width, based on estimation of channel conditions, without compromising BER constraints. To validate the energy savings, the energy consumption of basic operators has been obtained from real measurements for different bit-widths on a FPGA and a processor using soft SIMD. Results show that up to 62% of the dynamic energy consumption can be saved using this adaptive technique. The algorithms proposed for the low complexity selector used to choose the processing word-length at run time, without modifying the standard OFDM frame, are detailed in [38].

7.1.3. Optical Interconnections for 3D Multiprocessor Architectures

Participants: Jiating Luo, Pham Van Dung, Cédric Killian, Daniel Chillet, Olivier Sentieys.

To address the issue of interconnection bottleneck in multiprocessor on a single chip, we study how an Optical Network-on-Chip (ONoC) can leverage 3D technology by stacking a specific photonics die. The objectives of this study target: i) the definition of a generic architecture including both electrical and optical components, ii) the interface between electrical and optical domains, iii) the definition of strategies (communication protocol) to manage this communication medium, and iv) new techniques to manage and reduce the power consumption of optical communications. The first point is required to ensure that electrical and optical components can be used together to define a global architecture. Indeed, optical components are generally larger than electrical components, so a trade-off must be found between the size of optical and electrical parts. For example, if the need in terms of communications is high, several waveguides and wavelengths must be necessary, and can lead to an optical area larger than the footprint of a single processor. In this case, a solution is to connect (through the optical NoC) clusters of processors rather than each single processor. For the second point, we study how the interface can be designed to take applications needs into account. From the different possible interface designs, we extract a high-level performance model of optical communications from losses induced by all optical components to efficiently manage Laser parameters. Then, the third point concerns the definition of high-level mechanisms which can handle the allocation of the communication medium for each data transfer between tasks. This part consists in defining the protocol of wavelength allocation. Indeed, the optical wavelengths are a shared resource between all the electrical computing clusters and are allocated at run time according to application needs and quality of service. The last point concerns the definition of techniques allowing to reduce the power consumption of on-chip optical communications. The power of each Laser can be dynamically tuned in the optical/electrical interface at run time for a given targeted bit-error-rate. Due to the relatively high power consumption of such integrated Laser, we study how to define adequate policies able to adapt the laser power to the signal losses.

In [44], we proposed a wavelength reservation protocol handled by an Optical Network Interface (ONI) Manager for reconfigurable ONoC based on shared waveguide. It allows to efficiently allocate, at runtime, the optical communication channels for a manycore architecture. We described the ONI manager architecture and reservation protocol. Synthesis results in a 28nm FDSOI technology demonstrated that our interface can support a clock frequency up to 550 MHz with 6 wavelengths managed. From these results, we can be optimistic about the scaling of the ONoC and its capacity to manage a large number of processors and more wavelengths.

In [55], we explored the trade-off among channel bandwidth alternatives, performance, area and power. We showed that the channel size has a strong impact on the system performance and cost. We employed synthetic and real application traffic executed on the GEM5 simulator. As a result, we show that different channel bandwidths can improve the execution time of an application up to 75%, while including low area and power penalties.

7.1.4. Arithmetic Operators for Cryptography and Fault-Tolerance

Participants: Arnaud Tisserand, Emmanuel Casseau, Nicolas Veyrat-Charvillon, Karim Bigou, Franck Bucheron, Jérémie Métairie, Gabriel Gallin.

Arithmetic Operators for Fast and Secure Cryptography.

Our paper [36], presented at CHES, describes a new RNS modular multiplication algorithm for efficient implementations of ECC over $GF(p)$. Thanks to the proposition of RNS-friendly Mersenne-like primes, the proposed RNS algorithm requires 2 times less moduli than the state-of-art ones, leading to 4 times less precomputations and about 2 times less operations. FPGA implementations of our algorithm are presented, with area reduced up to 46 %, for a time overhead less than 10 %. Other RNS algorithms and implementations have been presented at RAIM [66].

Scalar recoding is popular to speed up ECC (elliptic curve cryptography) scalar multiplication: non-adjacent form, double-base number system, multi-base number system (MBNS). Ensuring uniform computation profiles is an efficient protection against some side channel attacks (SCA) in embedded systems. Typical ECC scalar multiplication methods use two point operations (addition and doubling) scheduled according to secret scalar digits. Euclidean addition chains (EAC) offer a natural SCA protection since only one point operation

is used. Computing short EACs is considered as a very costly operation and no hardware implementation has been reported yet. We designed an hardware recoding unit for short EACs which works concurrently to scalar multiplication. It has been integrated in an in-house ECC processor on various FPGAs. The implementation results show similar computation times compared to non-protected solutions, and faster ones compared to typical protected solutions (e. g. 18 % speed-up over 192 b Montgomery ladder). A paper [62] has been presented at Compas conference.

In a collaboration with University College Cork (Ireland), we worked on the design of secure multipliers for asymmetric cryptography using asynchronous circuits. A common paper has been published at ASYNC Conference [37]. In this paper, a specially adjusted Latch-less Asynchronous Charge Sharing Logic (LACSL) is developed to inherently defend such architecture against DPA attacks. The proposed logic provides input data independent low-power/energy consumption which is attributed to interleaved charge sharing stages with non-static elements involved in the data path. A 32-bit LACSL Montgomery Multiplier (case study) is extensively tested through HSPICE simulations and great consistency in power/energy consumption is achieved. The normalized energy deviation and normalized standard deviation are only 0.048 and 0.011, respectively. Compared with the original ACSL implementation, besides the impressive energy coherence, 42% energy saving is demonstrated plus that the leakage power is 3.5 times smaller. Furthermore, the scalability of the proposed multiplier is explored where 64-bit, 128-bit and 256-bit designs are implemented. Again, great energy consistency is found with the highest deviation being 0.5%.

In collaboration with D. Pamula, we worked on fast and secure finite field multipliers for $GF(2^m)$ arithmetic, a paper has been presented at DSD conference [53]. It presents details on fast and secure $GF(2^m)$ multipliers dedicated to elliptic curve cryptography applications. Presented design approach aims at high efficiency and security against side channel attacks of a hardware multiplier. The security concern in the design process of a $GF(2^m)$ multiplier is quite a novel concept. Basing on the results obtained in course of conducted research it is argued that, as well as efficiency of the multiplier impacts the efficiency of the cryptoprocessor, the security level of the multiplier impacts the security level of the whole cryptoprocessor. Thus the goal is to find a tradeoff, to compromise efficiency, in terms of speed and area, and security of the multiplier. We intend to secure the multiplier by masking the operation, either by uniformization or by randomization of the power consumption of the device during its work. The design methodology is half automated. The analyzed field sizes are the standard ones, which ensure that a cryptographic system is mathematically safe. The described architecture is based on principles of Mastrovito multiplication method. It is very flexible and enables to improve the resistance against side channel attacks without degrading the multiplier efficiency.

In a collaboration with G. Abozaid (EJUST University Egypt), we worked on the FPGA implementation of arithmetic operators for very large numbers (millions of bits) in fully homomorphic encryption (FHE) applications. A journal paper has been published in IEEE Embedded Systems Letters [18].

ECC Crypto-Processor with Protections Against SCA.

A dedicated processor for elliptic curve cryptography (ECC) is under development. Functional units for arithmetic operations in $GF(2^m)$ and $GF(p)$ finite fields and 160-600-bit operands have been developed for FPGA implementation. Several protection methods against side channel attacks (SCA) have been studied. The use of some number systems, especially very redundant ones, allows one to change the way some computations are performed and then their effects on side channel traces. This work is done in the PAVOIS project. An ASIC version of the processor is under development and should be sent for fabrication in the beginning of 2016.

A. Tisserand has been invited speaker at the conference on elliptic curve cryptography (ECC): "Hardware Accelerators for ECC and HECC" [29].

Arithmetic Operators and Crypto-Processor for HECC.

In the HAH project, we study and prototype efficient arithmetic algorithms for hyperelliptic curve cryptography for hardware implementations (on FPGA circuits). We study new advanced arithmetic algorithms and representations of numbers for efficient and secure implementations of HECC in hardware. First results have been published in Compas conference [60] and RAIM workshop [68].

Arithmetic Operators for Fault Tolerance.

In the ARDyT and Reliasic projects, we work on computation algorithms, representations of numbers and hardware implementations of arithmetic operators with integrated fault detection (and/or fault tolerance) capabilities. The target arithmetic operators are: adders, subtractors, multipliers (and variants of multiplications by constants, square, FMA, MAC), division, square-root, approximations of the elementary functions. We study two approaches: residue codes and specific bit-level coding in some redundant number systems for fault detection/tolerance integration at the arithmetic operator/unit level. FPGA prototypes are under development.

7.2. Compilation and Synthesis for Reconfigurable Platform

7.2.1. Adaptive dynamic compilation for low power embedded systems

Participants: Steven Derrien, Simon Rokicki.

Just-in-time (JIT) compilers have been introduced in the 1960s and became popular in the mid-1990s with the Java virtual machine. The use of JIT techniques for bytecode languages brings both portability and performance, making it an attractive solution for embedded systems, as evidenced by the Dalvik framework used by Android.

When targeting embedded systems, JIT compilation is even more challenging. First, because embedded systems are often based on architectures with an explicit use of Instruction- Level Parallelism (ILP), such as Very Long Instruction Word (VLIW) processors. Those architectures are highly dependent of the quality of the compilation, mainly because of the instruction scheduling phase performed by the compiler. The other challenge lies in the high constraints of the embedded system: the energy and execution time overhead due to the JIT compilation must be carefully kept under control. This is even more true if the JIT system is to be used in the context of a heterogeneous multi-core system with support dynamic task migration for heterogeneous ISA cores and/or support dynamically reconfigurable machines.

To address these challenges, we are currently studying how it is possible to take advantage of custom hardware to speed-up (and reduce the energy cost of) the JIT compilation stage. In this framework, basic optimizations and JIT management are performed in software, while the compilation back-end is implemented by means of specialized hardware. This back-end involves both instruction scheduling and register allocation, which are known to be the most time consuming stages of such a compiler. The first results are very encouraging, and we are finalizing an FPGA-based demonstration of the system.

7.2.2. Design Tools for Reconfigurable Video Coding

Participants: Emmanuel Casseau, Yaset Oliva.

In the field of multimedia coding, standardization recommendations are always evolving. To reduce design time taking benefit of available SW and HW designs, Reconfigurable Video Coding (RVC) standard allows defining new codec algorithms. The application is represented by a network of interconnected components (so called actors) defined in a modular library and the behaviour of each actor is described in the specific RVC-CAL language. Dataflow programming, such as RVC applications, express explicit parallelism within an application. However general purpose processors cannot cope with both high performance and low power consumption requirements embedded systems have to face. We have investigated the mapping of RVC applications onto a dedicated multiprocessor platform. Actually, our goal is to propose an automated co-design flow based on the RVC framework. The design flow starts with the Dynamic Dataflow and CAL descriptions of an application and goes up to the deployment of the system onto the hardware platform. We also propose a framework to explore dynamic mapping algorithms for multiprocessors systems. Such an algorithm should be capable of computing a more efficient workload repartition based on the current configuration and performances of the system. The targeted platform is composed of several Processing Elements (PE). They follow a hierarchical organization: one PE plays the role of master and the others are slaves. The master assigns tasks (actors) to the slaves. The slaves execute the application tasks. The system has been implemented on a Zynq platform. The mapping is computed at runtime on the ARM processor while two clusters of 8 Microblazes each play the role of slaves. The DDR memory is split into two sections: one is reserved to the

Master and the other one is shared with the slaves. This later contains the actor's code. On the FPGA, the Microblazes are connected to private memories through the Local Memory Bus (LMB) that store the runtime copy. A common shared memory is used for the data exchanges between the processors. It contains the FIFOs for token exchanges between actors. The dynamic mapping algorithm aims at increasing data throughput. It starts by gathering the performance metrics of the system. It then identifies the processor with the highest workload. The algorithm evaluates the gain when moving the actor to one of the other processors. The migration is only valuable if the overhead of moving the actor is less than the gain. The actor that would lead to the highest gain is selected for migration. As a use case, we implement an MPEG-4 decoder algorithm onto a multi-core heterogeneous system deployed onto the Zynq platform from Xilinx [61] [69]. This work is done in collaboration with Lab-STICC Lorient.

7.2.3. High-Level Synthesis Based Rapid Prototyping of Software Radio Waveforms

Participants: Emmanuel Casseau, Mai Thanh Tran.

Software Defined Radio (SDR) is now becoming a ubiquitous concept to describe and implement Physical Layers (PHYs) of wireless systems. In this context, FPGA (Field Programmable Gate Array) technology is expected to play a key role. To this aim, leveraging the nascent High-Level Synthesis (HLS) tools, a design flow from high-level specifications to Register-Transfer Level (RTL) description can be thought to generate processing blocks that can be reconfigured at run-time. We thus propose a methodology for the implementation of run-time reconfiguration in the context of FPGA-based SDR. The design flow allows the exploration between dynamic partial reconfiguration and control signal based multi-mode design. This architectural tradeoff relies upon HLS and its associated design optimizations. We apply the methodology to the architectural exploration of a Fast Fourier Transform (FFT) for Long Term Evolution (LTE) standard as a use case.

7.2.4. Optimization of loop kernels using software and memory information

Participant: Angeliki Kritikakou.

Compilers optimize the compilation sub-problems one after the other, following an order which leads to less efficient solutions because the different sub-problems are independently optimized taking into account only a part of the information available in the algorithms and the architecture. In [19], we have presented an approach which applies loop transformations in order to increase the performance of loop kernels. The proposed approach focuses on reducing the L1, L2 data cache and main memory accesses and the addressing instructions. Our approach exploits the software information, such as the array subscript equations, and the memory architecture, such as the memory sizes. Then, it applies source-to-source transformations taking as input the C code of the loop kernels and producing a new C code which is compiled by the target compiler. We have applied our approach to five well-known loop kernels for both embedded processors and general purpose processors. From the obtained experimental results we observed speedup gains from 2 up to 18. [21] presents a new methodology for computing the Dense Matrix Vector Multiplication, for both embedded (processors without SIMD unit) and general purpose processors (single and multi-core processors with SIMD unit). The proposed methodology fully exploits the combination of the software (e.g., data reuse) and hardware parameters (e.g., data cache associativity) which are considered simultaneously giving a smaller search space and high-quality solutions. The proposed methodology produces a different schedule for different values of the (i) number of the levels of data cache; (ii) data cache sizes; (iii) data cache associativities; (iv) data cache and main memory latencies; (v) data array layout of the matrix and (vi) number of cores. With our experimental results we show that the proposed approach achieves increased performance than ATLAS state-of-the-art library with a speedup from 1.2 up to 1.45.

7.2.5. Leveraging Power Spectral Density for Scalable System-Level Accuracy Evaluation

Participants: Benjamin Barrois, Olivier Sentieys.

The choice of fixed-point word-lengths critically impacts the system performance by impacting the quality of computation, its energy, speed and area. Making a good choice of fixed-point word-length generally requires solving an NP-hard problem by exploring a vast search space. Therefore, the entire fixed-point refinement process becomes critically dependent on evaluating the effects of accuracy degradation. In [34], a novel technique for the system-level evaluation of fixed-point systems, which is more scalable and that renders better accuracy, was proposed. This technique makes use of the information hidden in the power-spectral density of quantization noises. It is shown to be very effective in systems consisting of more than one frequency sensitive components. Compared to state-of-the-art hierarchical methods that are agnostic to the quantization noise spectrum, we show that the proposed approach is $5\times$ to $500\times$ more accurate on some representative signal processing kernels.

7.3. Interaction between Algorithms and Architectures

7.3.1. Sensor-Aided Non-Intrusive Load Monitoring

Participants: Xuan-Chien Le, Olivier Sentieys.

Non-Intrusive Load Monitoring (NILM) plays an important role in energy management and energy reduction in buildings and homes. An NILM system does not need a large amount of deployed power meters to monitor the power usage of home devices. Instead, only one meter on the main power line is necessary to detect and identify the operating devices. There are many approaches to solve the problem of device determination in NILM. The features applied in low-frequency based approach essentially include the step-change (or edge) and the steady state. In [47] we introduced three algorithms to solve the l_1 -norm minimization problem in NILM and results on power measurements obtained from a real appliance deployment. With a small number of devices, the obtained precision varies from 75% to 99%, depending on the tolerance criterion to determine the steady state of a given device.

7.3.2. Posture and Gesture Recognition using Wireless Body Sensor Networks

Participants: Arnaud Carer, Alexis Aulery, Olivier Sentieys.

The BoWi project (Body World Interactions) aims at designing a Wireless Body Sensor Network (WBSN) for accurate Gesture and Body Movement estimation with extremely severe constraints in terms of footprint and power consumption. Advantages of such system mainly come from its possible use in indoor or outdoor environments without any additional equipment. The 3D geolocation approach will combine radio communication distance measurement and inertial sensors and it will also strongly benefit from cooperative techniques based on multiple observations and distributed computation. Different types of applications, as health care, activity monitoring and environment control, are considered and evaluated along with a human-machine interface expertise.

In [32] we presented three different use cases of WBSN for posture and gesture recognition developed by increasing demands in terms of accuracy: posture recognition, gesture recognition and motion capture. This work is based on a simulator designed to explore algorithmic solutions for posture and gesture identification. Simulation results were performed with a set of different algorithm and sensor proposals for three usages including a Principal Component Analysis (PCA) for posture classification. We show how sensor and algorithm can be carefully chosen according to application scenarios while minimising implementation complexity.

For applications based on predefined postures such as environment control and physical rehabilitation, we show in [31] that low cost and fully distributed solutions, that minimize radio communications, can be efficiently implemented. Considering that radio links provide distance information, we also demonstrate that the matrix of estimated inter-node distances offers complementary information that allows for the reduction of communication load. Our results are based on a simulator that can handle various measured input data, different algorithms and various noise models. Simulation results are useful and used for the development of real-life prototype.

7.3.3. Energy Harvesting and Power Management

Participants: Olivier Sentieys, Arnaud Carer, Trong-Nhan Le.

To design autonomous Wireless Sensor Networks (WSNs) with a theoretical infinite lifetime, energy harvesting (EH) techniques have been recently considered as promising approaches. Ambient sources can provide everlasting additional energy for WSN nodes and exclude their dependence on battery.

In [24], an efficient energy harvesting system which is compatible with various environmental sources, such as light, heat, or wind energy, was proposed. Our platform takes advantage of double-level capacitors not only to prolong system lifetime but also to enable robust booting from the exhausting energy of the system. Simulations and experiments show that our multiple-energy-sources converter (MESCC) can achieve booting time in order of seconds. Although capacitors have virtual recharge cycles, they suffer higher leakage compared to rechargeable batteries. Increasing their size can decrease the system performance due to leakage energy. Therefore, an energy-neutral design framework providing a methodology to determine the minimum size of those storage devices satisfying energy-neutral operation (ENO) and maximizing system quality-of-service (QoS) in EH nodes, when using a given energy source, was also proposed. Experiments validating this framework are performed on a real WSN platform with both photovoltaic cells and thermal generators in an indoor environment. Moreover, simulations on OMNET++ showed that the energy storage optimized from our design framework is used up to 93.86%.

A Power Manager (PM) is usually embedded in EH wireless nodes to adapt the computation load by changing their wake-up interval according to the harvested energy. In order to prolong the network lifetime, the PM must ensure that every node satisfies the Energy Neutral Operation (ENO) condition. However, when a multi-hop network is considered, changing the wake-up interval regularly may cripple the synchronization among nodes and therefore, degrade the global system Quality of Service (QoS). In [25], a Wake-up Variation Reduction Power Manager (WVR-PM) was proposed to solve this issue. This PM is applied for wireless nodes powered by a periodic energy source (e.g. light energy in an office) over a constant cycle of 24 hours. Not only following the ENO condition, our power manager also reduces the wake-up interval variations of WSN nodes. Based on this PM, an energy-efficient protocol, named Synchronized Wake-up Interval MAC (SyWiM), was also proposed. OMNET++ simulation results using three different harvested profiles show that the data rate of a WSN node can be increased up to 65% and the latency reduced down to 57% compared to state-of-the-art PMs. Validations on a real WSN platform have also been performed and confirmed the efficiency of our approach.

7.3.4. Signal Processing for High-Rate Optical Communications

Participants: Trung-Hien Nguyen, Olivier Sentieys, Arnaud Carer.

M-ary quadrature amplitude modulation (m -QAM) combined with coherent detection and digital signal processing (DSP) is a promising candidate for the implementation of next generation optical transmission systems. However, as the number of modulation levels increases, the sensitivity to system imperfections such as phase noise of the transmitter and the local oscillator lasers or fiber nonlinearities is exacerbated. Moreover, the amplitude and phase imbalances between the in-phase (I) and quadrature (Q) channels in the transmitter (Tx) and the front-end of the receiver (Rx), which is often referred to as IQ imbalance, is also troublesome if not compensated

In [52], we proposed a novel simple blind adaptive IQ imbalance compensation based on a decision-directed least-mean-square (DD-LMS) algorithm integrated to a modified butterfly FIR filter configuration. Since only 2 FIR filter coefficients-sets are used, instead of 4 in the conventional configuration, the time for updating the coefficients and the hardware resources (such as coefficient memories and number of look-up tables) in real time field-programmable gate array (FPGA) platforms is then reduced using this method. A reduction in hardware complexity by a factor of about 3 is achieved by the proposed joint method. The proposed structure is experimentally validated with a 40Gbit/s 16-QAM signal. A 7dB power penalty reduction is experimentally achieved at a bit error rate (BER) of 10^{-3} in the presence of a 10 degree phase imbalance, confirming the effectiveness of the proposed algorithm. The equalization capability remains even in the presence of group velocity dispersion along the link, which is numerically confirmed with optical fiber transmission up to 1200 km and 20 phase imbalance.

In [50], circular harmonic expansion-based carrier frequency offset estimation was investigated for optical m -QAM communication systems. The proposed method, combined with a gradient-descent algorithm, shows better performance compared to already proposed VVMFOE and 4th-power methods.

CAMUS Team

7. New Results

7.1. Formal Proofs for an Ordering Relation in Explicitly Parallel Programs

Participants: Alain Ketterlin, Éric Violard.

This project is a collaborative work with the COMPSYS Inria Team, in Lyon. Participants are: Paul Feautrier, Tomofumi Yuki.

The growing need to make use of available parallelism has led to new explicitly parallel language constructs. These constructs are usually grouped under the term *Task Parallelism*, because they aim to go beyond “simple” *Data Parallelism* (i.e., loop and array-based parallelism). Prominent examples of languages integrating task parallelism are X10 (<http://x10-lang.org>) and variants, Cilk (<http://supertech.csail.mit.edu/cilk/>), and recent versions of OpenMP (<http://www.openmp.org>). Most of the work on such languages has focused on efficient run-time support for *tasks*, in contrast with *threads*, i.e., for programs generating potentially large numbers of distinct tasks with explicit (but arbitrary) ordering between the tasks. However, little attention has been given to the static analysis and optimization of explicitly parallel programs, probably because their properties are much harder to formalize, compared to their sequential counterpart. Starting with the work of our colleagues Paul Feautrier and Tomofumi Yuki, from the Compsys team in Lyon, we have advanced the formalization and formally proved several properties of some fundamental building blocks for the analysis of certain classes of explicitly parallel programs.

Task parallelism is usually based on a few syntactic constructs to represent tasks and their synchronization. We use X10’s terminology (and syntax, with simplifications), but the corresponding constructs of other languages is usually obvious. Across all languages one finds a construct to start (or *spawn*) an asynchronous task, named `async` in X10, and a “container” construct, named `finish` in X10, whose role is to wait for the completion of all task spawned during the execution of its body. Given that these constructs allow the parallel execution of pieces of the program, a first question arises: is there a static (i.e., compile-time) way to decide whether two given statements are ordered, i.e., that the first necessarily executes before the other. Feautrier and Yuki (with colleagues) have defined such a criterion for programs made of `async` and `finish` [33], along with arbitrary statements and for-loops, defining the so-called *polyhedral fragment* of X10. The resulting (partial) relation, called *happens-before*, opens the door to various static analyses, like data-dependence analysis, which are at the heart of a range of optimization techniques. Here is a quick example:

```

finish
for i in ...
  async
  for j in ...
    S(i,j)

```

$S(i, j)$ happens before $S(i', j')$ iff $i = i' \wedge j < j'$

The resulting condition, $i = i' \wedge j < j'$, defines exactly the situation in which two statement executions are ordered, and can be seen as an appropriate extension of the lexicographic order to explicitly parallel programs.

Our work on this basis has been to take the formal definition of happens-before (HB), and implement it in Coq (<https://coq.inria.fr>). The goal was first to prove various properties of the relation, like transitivity, and second to provide a formal proof of both correctness and completeness of HB itself. The first part has been fairly immediate, due to the high representative power of Coq. The second part took more time, and involved several new contributions. The major part of the work went into defining a formal semantics for the fragment of X10 needed by the definition of HB. Given the semantics, it was possible to obtain the relation between a program and its trace(s), and then to prove that HB is correct (i.e., if HB states that one statement executes before another, then these statements appear in order in all possible traces of the program), and that HB is complete (i.e., that statements that are always ordered in traces are actually recognized as such by HB). The complete proof scripts are available on the Inria forge ([gforge.inria.fr](https://forge.inria.fr)), under the `x10-coq` project.

Further work has also started on extending *happens-before* to X10 programs using synchronization primitives called *clocks*, which are basically *barriers*, where distinct tasks can wait for each other. Since an unrestricted use of synchronization barriers can lead to deadlocks, X10 introduces “implicit clocks”, which are introduced (and scoped) by a `finish` construct, on which a task can “register”, and whose scoping rules ensure that any program point can only use the single “nearest” clock. These restrictions offer termination guarantees, which in turn enables a sound *happens-before* relation between statement instances. The “clock-less” HB relation can then be modified to take into account the additional ordering imposed by clocks. We have started work to update the semantics to the case of implicit clocks, and to formalize this extension in Coq.

7.2. Validity Conditions for Transformations of Non-Affine Programs

Participants: Alain Ketterlin, Philippe Clauss.

This project is a collaborative work with the CORSE Inria Team, in Grenoble. Participant is: Fabrice Rastello.

Representing loop nests with the help of the polyhedral model has been a powerful and fruitful strategy to enable automatic optimization and parallelization. However, this model places strong requirements on the input program, and in many cases these requirements are hard to meet. Because they are based on linear programming, polyhedral techniques require every constraint to be affine in loop counters and parameters. While this is easily verified for loop bounds in a large majority of programs, the same constraint imposed to memory access functions is often too strong. There are several reasons for this. First, programmers often linearize multi-dimensional arrays, turning straightforward accesses like `t[i][j]` into `t1[i*n+j]`, with the unfortunate effect of placing their program outside the scope of the polyhedral model. Second, optimization often happens late in the compilation process (or even during just-in-time compilation at run-time), where multi-dimensional array accesses have been transformed by the compiler itself, for the needs of its earlier passes. Third, complex data storage strategies for certain classes of arrays, e.g., band or triangular matrices, may introduce non-linear access functions, and this non-linearity must be taken into account, e.g., for locality optimization. And fourth, some access functions are almost completely unspecified, like in the case of indirect accesses (`t[s[i]]`) or abstract mappings (`t[f(i)]`).

Our goal is to extend polyhedral analysis techniques to cover at least some of these cases, and see how far we can push the limits of the fundamental algorithms beyond pure linearity. We have started by considering the case of multi-dimensional array linearization, where the code doesn’t provide access functions for all (original) dimensions, but rather a single access function, which is linear in loop counters but contains parametric coefficients. Here is an example illustrating our initial target, which is taken from the `gemver` program in the `polybench` suite:

```

for (i = 0; i < n; i++)
  for (j = 0; j < n; j++)
    // Was: A[i][j] = A[i][j] + u1[i] * v1[j] + ...;
S1:  *(n*i+A+j) = *(n*i+A+j) + *(u1+i) * *(v1+j) + ...;
for (i = 0; i < n; i++)
  for (j = 0; j < n; j++)
    // Was: x[i] = x[i] + beta * A[j][i] * y[j];

```

```
S2:  *(x+i) = *(x+i) + beta * *(n*j+A+i) * *(y+j);
// ...
```

The original form of the statements appear in comments, but what finally reaches the compiler is much more convoluted: basically, every array access appears as a pointer access whose effective address is a polynomial function mixing counters (i, j), array base addresses ($A, u1, v1, x, y$), and size parameters (n). In some other cases, the arrays have been “locally” linearized, i.e., the code still displays different arrays, but their inner dimensions have been linearized. In our example, statement S1 would appear as:

```
// Was: A[i][j] = A[i][j] + u1[i] * v1[j] + ...;
S1:  A[n*i+j] = A[n*i+j] + u1[i] + v1[j] + ...;
```

This is an important special case in practice, and its particular structure helps a lot, for example, when data dependence analysis is needed.

Extending current polyhedral techniques to deal with non-affine accesses is a formidable endeavor, requiring the adaptation of the many algorithms developed over decades for analysis, scheduling, and code generation. Rather, we have started by studying a specific task, with immediate practical impact: given a non-affine loop nest *and* a specific desired transformation, what are the conditions under which this condition is valid? It is not unreasonable to expect the transformation to be provided by other means than pure analysis, for instance to be suggested by profiling data. In this case, the problem we are left with is the one of testing whether the given transformation is valid. This in turn requires testing the emptiness of a “problematic system”. For any given loop nest, this can be written as:

$$\bigvee_{(A,A')} \exists(v,v') \text{s.t.}$$

$$\begin{aligned} & v \in \mathcal{D}_A \wedge v' \in \mathcal{D}_{A'} && \text{(domain)} \\ & \wedge v \prec_{lex} v' && \text{(originalschedule)} \\ & \wedge A(v) = A'(v) && \text{(sameaccesslocation)} \\ & \wedge T_A(v) \neg \prec_{lex} T_{A'}(v') && \text{(transformedschedule)} \end{aligned}$$

where A and A' range over pairs of potentially conflicting accesses, v and v' are iteration vectors, \mathcal{D}_A and $\mathcal{D}_{A'}$ are iteration domains, $A(v)$ and $A'(v')$ are access functions, and T_A and $T_{A'}$ are schedules. The condition under which the transformation is valid is the projection of this set on parameter dimensions, i.e., the elimination of all variables representing counters. The difficulty of this comes from the non-affine condition expressing the equality of access functions.

Building on previous work, we have devised a projection procedure that eliminates all counters and leaves a (usually complex) condition on parameters. We have also developed several simplification strategies, applied during elimination and also on the final result, that overall produces a test deciding whether the targeted transformation can be applied. For instance, on the fully linearized version of the previous examples, when deciding whether the following transformation is legal:

$$T_{S1}(0, i, j) = (0, i, j) \quad T_{S2}(1, i, j) = (0, j, i)$$

i.e., interchanging the second loop (around S2) and then applying fusion on both depth-2 loops, our elimination and simplification procedure produces the following run-time test:

```
if ( ((y+n >= x+2) && (x+n >= y+2))
    || ((n >= 2) && (n*n+A >= x+1) && (x >= A+1))
    || ((n >= 2) && (u1+n >= x+1) && (x+n >= u1+2))
    || ((n >= 2) && (n+v1 >= x+1) && (x+n >= v1+1))
    || ((n*n+A >= y+1) && (y >= A+1) && (n >= 2)) || ...)
    // Transformation invalid: run the original version...
else
    // Transformation valid: run the transformed version...
```

The reader may want to verify that this test actually corresponds to verifying that “arrays” do not overlap, but only as far as the given transformation requires it.

A systematic evaluation of our procedure on a benchmark suite has shown that the resulting tests are both accurate and incur very little run-time overhead. The overall mechanism compares favorably with alternative techniques aiming at dealing with non-affine access functions, which consist in statically reconstructing array dimensions [30]. This part of our work is ready for publication. However, to be completely competitive with alternative approaches, we need to find ways to complete the polyhedral compilation chain, with a prior effective scheduling algorithm and a posterior code generation algorithm.

7.3. Automatic Parallelization of Nonlinear Loops

Participants: Aravind Sukumaran-Rajam, Philippe Clauss.

Runtime code optimization and speculative execution are becoming increasingly prominent to leverage performance in the current multi- and many-core era. However, a wider and more efficient use of such techniques is mainly hampered by the prohibitive time overhead induced by centralized data race detection, dynamic code behavior modeling, and code generation. Most of the existing Thread Level Speculation (TLS) systems rely on naively slicing the target loops into chunks and trying to execute the chunks in parallel with the help of a centralized performance-penalizing verification module that takes care of data races. Due to the lack of a data dependence model, these speculative systems are not capable of doing advanced transformations, and, more importantly, the chances of rollback are high. The polyhedral model is a well-known mathematical model to analyze and optimize loop nests. The current state-of-art tools limit the application of the polyhedral model to static control codes. Thus, none of these tools can generally handle codes with while loops, indirect memory accesses, or pointers. Apollo (Automatic POLYhedral Loop Optimizer) is a framework that goes one step beyond and applies the polyhedral model dynamically by using TLS. Apollo can predict, at runtime, whether the codes are behaving linearly or not, and it applies polyhedral transformations on-the-fly.

Apollo has been extended to handle codes whose memory accesses and loop bounds are not necessarily linear [23], [14]. The proposed extension consists of modeling memory addresses that are accessed either as “tubes” obtained through linear regression, or as ranges. More generally, this approach expands the applicability of the polyhedral model at runtime to a wider class of codes. Plugging together both linear and nonlinear accesses to the dependence prediction model enables the application of polyhedral loop optimizing transformations even for nonlinear code kernels while also allowing a low-cost speculation verification.

This work takes part of Aravind Sukumaran-Rajam’s PhD thesis that has been defended November the 5th, 2015 [13].

7.4. Dynamic Code Generation for Speculative Polyhedral Optimization

Participants: Juan Manuel Martinez Caamano, Philippe Clauss.

We have developed a new runtime code generation technique for speculative loop optimization and parallelization, that allows to generate on-the-fly codes resulting from any polyhedral optimizing transformation of loop nests, such as tiling, skewing, loop fission, loop fusion or loop interchange, without introducing a penalizing time overhead. The proposed strategy is based on the generation of code bones at compile-time, which are parametrized code snippets either dedicated to speculation management or to computations of the original target program. These code bones are then instantiated and assembled at runtime to constitute the speculatively-optimized code, as soon as an optimizing polyhedral transformation has been determined. Their granularity threshold is sufficient to apply any polyhedral transformation, while still enabling fast runtime code generation. This strategy has been implemented in the speculative loop parallelizing framework Apollo.

7.5. The XFOR Programming Structure

Participants: Imen Fassi, Philippe Clauss, Cédric Bastoul.

We have proposed a new programming control structure called “xfor” or “multifor”, providing users a way to schedule explicitly the statements of a loop nest, and take advantage of optimization and parallelization opportunities that are not easily attainable using the standard programming structures, or using automatic optimizing compilers [19]. This is the PhD work of Imen Fassi, who started her work in 2013 and who defended her thesis November the 27th, 2015 [12].

It has been shown that xfor programs often reach better performance than programs optimized by fully automatic polyhedral compilers like Pluto [29]. It has also been shown that different versions of codes may perform very differently, although their memory behaviors are very similar. By analyzing further the origins of such performance differences, we noticed five important gaps in the currently adopted and well-established code optimization strategies [18], [19]: insufficient data locality optimization, excess of conditional branches in the generated code, too verbose code with too many machine instructions, data locality optimization resulting in processor stalls, and finally missed vectorization opportunities.

To ease and extend the usage of the XFOR structure, we have developed:

- Xfor-Wizard, which is a programming environment for XFOR programs, assisting users in writing XFOR codes and applying optimizing transformations. Automatic dependence analysis and comparisons against a referential code (XFOR-loops or classic for-loops) are achieved in order to help the user in ensuring semantic correctness of the written code.
- XFORGEN, which is a tool to automatically generate an XFOR code that is equivalent to for-loops that have been automatically transformed using a static polyhedral compiler. The generated XFOR code exhibits the parameters of the transformations that have been applied and thus can be modified for further optimizations.

7.6. Dynamic Optimization of Binary Code

Participants: Philippe Clauss, Alain Ketterlin.

This project is a collaborative work with the ALF Inria Team, in Rennes. Participants are: Erven Rohou and Nabil Hallou.

Automatic code optimizations have traditionally focused on source-to-source transformation tools and compiler IR-level techniques. Sophisticated techniques have been developed for some classes of programs, and rapid progress is made in the field. However, there is a persistent hiatus between software vendors having to distribute generic programs, and end-users running them on a variety of hardware platforms, with varying levels of optimization opportunities. The next decade may well see an increasing variety of hardware, as it has already started to appear particularly in the embedded systems market. At the same time, one can expect more and more architecture-specific automatic optimization techniques.

Unfortunately, many “old” executables are still being used although they have been originally compiled for now outdated processor chips. Several reasons contribute to this situation:

- commercial software is typically sold without source code (hence no possibility to recompile) and targets slightly old hardware to guarantee a large base of compatible machines;
- though not commercial, the same applies to most Linux distributions⁰ – for example Fedora 16 (released Nov 2011) is supported by Pentium III (May 1999)⁰;
- with the widespread cloud computing and compute servers, users have no guarantee as to where their code runs, forcing them to target the oldest compatible hardware in the pool of available machines.

All this argues in favor of binary-to-binary optimizing transformations. Such transformations can be applied either statically, i.e., before executing the target code, or dynamically, i.e., while the target code is running.

⁰with the exception of Gentoo that recompiles every installed package

⁰http://docs.fedoraproject.org/en-US/Fedora/16/html/Release_Notes/sect-Release_Notes-Welcome_to_Fedora_16.html

Dynamic optimization is mostly addressing adaptability to various architectures and execution environments. If practical, dynamic optimization should be preferred because it eliminates several difficulties associated with static optimization. For instance, when deploying an application in the cloud, the executable file may be handled by various processor architectures providing varying levels of optimization opportunities. Providing numerous different adapted binary versions cannot be a general solution. Another point is related to interactions between applications running simultaneously on shared hardware, where adaptation may be required to adjust to the varying availability of the resources. Finally, most code optimizations have a basic cost that has to be recouped by the gain they provide. Depending on the input data processed by the target code, an optimizing transformation may or may not be profitable.

We distinguish two classes of binary transformations:

1. code transformations that can be handled directly by analyzing and modifying the original binary code. We call such transformations *low-level binary transformations*;
2. code transformations that require a higher level of abstraction of the code in order to generate a very different, but semantically equivalent, optimized code. We call such transformations *high-level binary transformations*.

While we target both classes of transformations, the first was addressed by focusing on SSE to AVX transformations of vectorized codes [20].

In this work, we focus on SIMD ISA extensions, and in particular on the x86 SSE and AVX capabilities. Compared to SSE, AVX provides wider registers, new instructions, and new addressing formats. AVX has been first supported in 2011 by the Intel Sandy Bridge and AMD Bulldozer architectures. However, most existing applications take advantage only of SSE and miss significant opportunities. We show that it is possible to automatically convert SSE to AVX whenever profitable. The key characteristics of our approach are:

- we apply the transformation at run-time, i.e. when the hardware is known;
- we only transform hot loops (detected through very lightweight profiling), thus minimizing the overhead;
- we do *not* implement a vectorization algorithm in a dynamic optimizer, instead we recognize already statically vectorized loops, and convert them to a more powerful ISA at low cost.

For high-level binary transformations, we also focus on hot loops and loop nests appearing in executable codes. There is an important literature addressing automatic loop optimization and parallelization techniques. Such optimizations include combinations of loop interchange, loop fusion and fission, loop skewing, loop shifting and loop tiling. However, they are mostly applied at compile-time, either on the source code, or on an intermediate representation form of the code. The most advanced techniques are related to the polyhedral model.

Applying such advanced loop optimizing transformations at runtime, on a currently running binary code, without any previous knowledge, is our challenging goal. The same goal has been addressed in [8], but not at runtime. In this work, the binary code is analyzed and transformed without any constraint regarding the related time overhead. Candidate loops are identified regarding their compliance to the polyhedral model: the loop bounds and memory references must be convertible into linear functions of the loop indices. Then, compliant loop nests are translated into an equivalent program in C source code, in order to be used as input for the source-to-source polyhedral compiler Pluto [29]. The resulting optimized code is then compiled and re-injected into the original binary code.

While a similar approach should be considered to reach the same goal at runtime, it must be handled differently regarding three main issues:

1. At runtime, the time overhead of the employed analysis and optimization techniques must be small. Thus, any translation to source code, that would require costly steps for the de-compilation/re-compilation phases, must be avoided.
2. Static approaches, as the one presented in [8], can only handle loops that are syntactically compliant with the polyhedral model. However, it has been shown, with the Apollo framework, that loops

may exhibit a compliant behavior at runtime. Since we target runtime optimizations, we also can take advantage of the information that is only available at runtime, and maybe also use speculative techniques.

3. Binary codes may hide some interesting properties of the embedded loops, and may need very complex analysis techniques for discovering such properties. In short, a whole compiler for binary codes would be required.

To address these issues, we are currently investigating the strategy consisting first of translating, at runtime, any selected loop nest into the LLVM⁰ intermediate representation form (LLVM-IR). This representation offers several advantages:

- Analysis and transformation passes of the LLVM compiler can be used on-the-fly, in order to discover and compute relevant information, and to safely transform the code;
- The LLVM just-in-time compiler can be used to compile the optimized code, which is in LLVM-IR, as an executable;
- Existing tools for loop optimization can be used, as Polly⁰, for static polyhedral-compliant loops, or Apollo, for dynamic polyhedral-compliant loops.

Hence, this strategy requires a fast binary-to-LLVM-IR translator. For this purpose, we are currently using and extending McSema⁰, which is a library for translating the semantics of native code to LLVM-IR. McSema supports translation of x86 machine code, including integer, floating point, and SSE instructions. Control flow recovery is separated from translation, permitting the use of custom control flow recovery front-ends.

For McSema to be able to handle mostly any code, we had to parametrize carefully its translation rules, and also to add some x86 SSE instructions that were not handled. McSema was recently plugged to the Padrone platform. Thus, any hot loop nest is now automatically converted into LLVM-IR, as illustrated in Figure 2.

Instead of taking as input a binary file, McSema takes as input a code extract containing a hot loop nest, thanks to the code address provided by Padrone. Then, McSema builds the control flow graph of the input code and generates a corresponding LLVM-IR. The next step is to plug the polyhedral LLVM compiler Polly (phases *Canonicalization* to *CodeGeneration* in Figure 2), in order to generate automatically an optimized version of the target loop nest, that will be then compiled using the LLVM just-in-time compiler and re-injected in the running code.

7.7. Combining Locking and Data Management Interfaces

Participants: Jens Gustedt, Mariem Saied, Daniel Salas.

Handling data consistency in parallel and distributed settings is a challenging task, in particular if we want to allow for an easy to handle asynchronism between tasks. Our publication [5] shows how to produce deadlock-free iterative programs that implement strong overlapping between communication, IO and computation. The collaboration with Soumeya Hernane has continued after her thesis defence in 2013. It extends distributed lock mechanisms and combines them with implicit data management, and resulted in a journal submission, see [26].

A new implementation (ORWL) of our ideas of combining control and data management in C has been undertaken, see 6.9. In previous work it has demonstrated its efficiency for a large variety of platforms. In 2015, work on the ORWL model and library has gained vigor with the thesis of Mariem Saied (Inria) and Daniel Salas (INSERM). We also now collaborate on that subject with the TADAAM project team from Inria Bordeaux, where a postdoc has been hired through Inria funding.

In 2015, a new domain specific language (DSL) has been developed that largely eases the implementation of applications with ORWL. In its first version it provides an interface for stencil codes, but extensions towards other types of applications are on their way. In addition, work has been started to encapsulate imaging applications that use certain pipeline patterns to describe dependencies between computational task.

⁰<http://llvm.org>

⁰<http://polly.llvm.org>

⁰<https://github.com/trailofbits/mcsema>

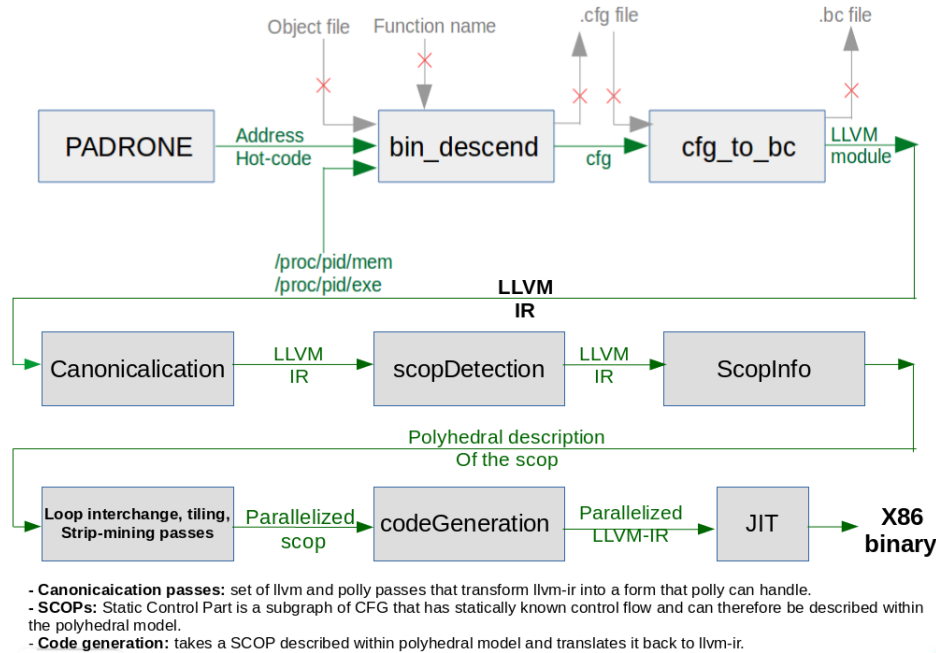


Figure 2. High-level Binary Loop Optimization through LLVM-IR

7.8. Efficient Execution of Polyhedral Codes on GPU and CPU+GPU Systems

Participants: Jean-François Dollinger, Vincent Loechner.

This is the main result of Jean-François Dollinger's PhD, started in 2012 and defended on July the 1st, 2015 [11].

Recent architectures complexity makes it difficult to statically predict the performance of a program. We have developed a reliable and accurate parallel loop nests execution time prediction method on GPUs for polyhedral codes. It is entirely automatic, and it based on three stages: static code generation, offline profiling on the target architecture, and online prediction.

In addition, we derived two techniques to fully exploit the computing resources at disposal on a computer. The first technique consists in jointly using all CPU cores and GPUs for executing a code. In order to achieve good performance, it is mandatory to consider load balance, in particular by predicting the execution time of a loop nest distributed on all those processing units. The runtime scheduler uses the profiling results to predict the execution times and adjust the parallel loop bounds to ensure load balance. The second technique puts CPU and GPU in a competition: instances of the considered code are simultaneously executed on CPU and GPU. The winner of the competition notifies its completion to the other instance, implying its termination.

7.9. Interactive Code Restructuring

Participants: Cédric Bastoul, Oleksandr Zinenko, Stéphane Huot.

This work falls within the exploration and development of semi-automatic programs optimization techniques. It consists in designing and evaluating new visualization and interaction techniques for code restructuring, by defining and taking advantage of visual representations of the underlying mathematical model. The main goal is to assist programmers during program optimization tasks in a safe and efficient way, even if they neither

have expertise into code restructuring nor knowledge of the underlying theories. This project is an important step for the efficient use and wider acceptance of semi-automatic optimization techniques, which are still tedious to use and incomprehensible for most programmers. More generally, this research is also investigating new presentation and manipulation techniques for code, algorithms and programs, which could lead to many practical applications: collaboration, tracking and verification of changes, visual search in large amount of code, teaching, etc.

This is a rather new research direction which strengthens CAMUS's static parallelization and optimization issue. It is a joint work with two Inria teams specialized in interaction: EX-SITU at Inria Saclay (contact: Oleksandr Zinenko) and MJOLNIR at Inria Lille (contact: Stéphane Huot).

In 2015, we presented our interactive tool, *Clint*, that maps direct manipulation of the visual representation to polyhedral program transformations with real-time semantics preservation feedback. We conducted two user studies showing that *Clint*'s visualization can be accurately understood by both experts and non-expert programmers, and that the parallelism can be extracted better from *Clint*'s representation than from the source code in many cases [21]. We are planing a first release of that tool in the coming year.

7.10. Automatic Generation of Adaptive Simulation Codes

Participants: Cédric Bastoul, César Sabater.

Compiler automatic optimization and parallelization techniques are well suited for some classes of simulation or signal processing applications, however they usually don't take into account neither domain-specific knowledge nor the possibility to change or to remove some computations to achieve "good enough" results. Quite differently, production simulation and signal processing codes have adaptive capabilities: they are designed to compute precise results only where it matters if the complete problem is not tractable or if the computation time must be short. In this research, we design a new way to provide adaptive capabilities to compute-intensive codes automatically, inspired by Adaptive Mesh Refinement a classical numerical analysis technique to achieve precise computation only in pertinent areas. It relies on domain-specific knowledge provided through special pragmas by the programmer in the input code and on polyhedral compilation techniques, to continuously regenerate at runtime a code that performs heavy computations only where it matters at every moment. A case study on a fluid simulation application shows that our strategy enables dramatic computation savings in the optimized portion of the application while maintaining good precision, with a minimal effort from the programmer.

This research direction started in 2015 and complements our other efforts on dynamic optimization. We are in the process of a collaboration with Inria Nancy Grand Est team TONUS, specialized on applied mathematics (contact: Philippe Helluy), to bring models and techniques from this field to compilers. First results, investigated during the Inria Internship Program of César Sabater, have been presented to the SimRace international conference dedicated on industrial fluid simulation applications [16].

7.11. Polyhedral Compiler White-Boxing

Participants: Cédric Bastoul, Lénaïc Bagnères, Oleksandr Zinenko, Stéphane Huot.

While compilers offer a fair trade-off between productivity and executable performance in single-threaded execution, their optimizations remain fragile when addressing compute-intensive code for parallel architectures with deep memory hierarchies. Moreover, these optimizations operate as black boxes, impenetrable for the user, leaving them with no alternative to time-consuming and error-prone manual optimization in cases where an imprecise cost model or a weak analysis resulted in a bad optimization decision. To address this issue, we researched and designed a technique allowing to automatically translate an arbitrary polyhedral optimization, used internally by loop-level optimization frameworks of several modern compilers, into a sequence of comprehensible syntactic transformations as long as this optimization focuses on scheduling loop iterations. With our approach, we open the black box of the polyhedral frameworks enabling users to examine, refine, replay and even design complex optimizations semi-automatically in partnership with the compiler.

This research started in 2014 and we found the first solution in 2015. It has been conducted as a joint work between teams in compiler technologies (CAMUS and Inria Saclay's POSTALE team) and teams in interaction (EX-SITU at Inria Saclay and MJOLNIR at Inria Lille). The first paper on this has been accepted in 2015 to be presented in one of the top conferences on optimization techniques: CGO 2016 [15]. Subsequent work and a first release of the tool implementing the technique is planned during 2016.

CAMEL Project-Team

7. New Results

7.1. The Logjam attack against the discrete logarithm

Participants: Pierrick Gaudry, Emmanuel Thomé [contact], Paul Zimmermann.

Together with colleagues from the Prosecco project-team and with other colleagues, we exhibited a new attack against the TLS protocol when using discrete logarithms [15]. A proof-of-concept of the attack was demonstrated using the CADO-NFS software. This paper obtained the best paper award at the ACM CCS 2015 conference, and received significant media coverage both in the specialized and non-specialized press.

7.2. Other results related to discrete logarithm

Participant: Pierrick Gaudry [contact].

Our 2014 work [16], in collaboration with Barbulescu, Guillevic and Morain, improving the practical aspects of discrete logarithm computation in quadratic extensions and reducing the theoretical complexity in the “medium characteristic case” has been published in Eurocrypt 2015.

In collaboration with Barbulescu and Kleinjung we have proposed in [17] to revisit an old construction of Schirokauer for discrete logarithms in extension fields. It is well suited for problems coming from pairings where the primes often have a special form.

With Galbraith we wrote a survey about the discrete logarithm problem in the context of elliptic curves [13].

7.3. Fast arithmetic for faster integer multiplication

Participants: Svyatoslav Covanov [contact], Emmanuel Thomé.

The paper [20] describes an algorithm for the multiplication of two n -bit integers. It achieves the best asymptotic complexity bound $O(n \log n \cdot 4^{\log^* n})$ under a hypothesis on the distribution of generalized Fermat primes of the form $r^{2^k} + 1$. This hypothesis states that there always exists a sufficiently small interval in which we can find such a prime. Experimental results give evidence in favor of this assumption. A journal submission is planned shortly.

7.4. Certificates for exact linear algebra computations

Participant: Emmanuel Thomé [contact].

The paper [21], in collaboration with Jean-Guillaume Dumas and Erich Kaltofen, is a preliminary version of a research work that has then been pursued, and that solves an open question of proving the correctness of some specific linear algebra computations. It emerged from practical techniques which had been used for this purpose for a while, and for which improvements were obtained. Submission plans for this work are yet to be finalized.

7.5. Computing Jacobi’s theta function in quasi-linear time

Participant: Hugo Labrande [contact].

We designed a new algorithm that improves the complexity of computing the value of the Jacobi theta function, $\theta(z, \tau)$ to arbitrary precision [23]. The algorithm uses a quadratically convergent sequence similar to the complex AGM, as well as Newton’s method; its complexity is $O(\mathcal{M}(n) \log n)$ for computing the value up to an error bounded by 2^{-n} , which is an improvement over the state-of-the-art complexity of $O(\mathcal{M}(n)\sqrt{n})$. Here, $\mathcal{M}(n)$ denotes the time taken by a multiplication of two n -bit numbers. We provide bounds on the loss of significant digits incurred during the computation. The algorithm was implemented using GNU MPC, showing practical improvement over (our optimized implementation of) existing algorithms for precision above approximately 300,000 bits. The paper was submitted to *Mathematics of Computation*.

7.6. Construction of sparse polynomial systems with many positive solutions

Participant: Pierre-Jean Spaenlehauer [contact].

In collaboration with Frédéric Bihan (Univ. Savoie Mont-Blanc), we propose a variant of the classical Viro method to construct polynomial systems with prescribed monomial support and many solutions whose coordinates are all positive [19]. This is an asymptotic construction which has strong connections with tropical and convex geometry, and which involves computational problems such as low-rank matrix completion.

7.7. Small certificates of inconsistency of quadratic fewnomial systems

Participant: Pierre-Jean Spaenlehauer [contact].

In collaboration with Jean-Charles Faugère (EPI PolSys) and Jules Svartz (Min. de Éducation Nationale), we studied the problem of certifying the inconsistency of sparse quadratic polynomial systems. Finding certificates of inconsistency is a classical problem in computational commutative algebra, and these certificates are in general of size exponential in the input size. We identify families of quadratic fewnomial systems for which there exist certificates of size linear in the size of the input and we propose algorithms to compute them in polynomial time.

7.8. Cracking passphrases based on famous sentences

Participant: Hugo Labrande [contact].

We proposed a method to attack passwords based on famous sentences, which are rather widespread [18]: we showed a method to construct large dictionaries using only publicly-available sources (e.g. Wikipedia) and modest computing power. The resulting dictionaries were able to crack millions of passphrases, among which a 55-character long one, and some that do not appear to have been cracked before. Our work thus shows that using famous sentences as passwords is not secure at all, as any attacker, even those with low skills and very modest computational resources, can guess them.

CARTE Project-Team

7. New Results

7.1. Computability and Complexity

- **Complexity of stream functions and higher-order complexity.** We have pursued our works on higher-order complexity and the complexity of stream functions. Both notions are closely related as any function from natural numbers to natural numbers can be seen as a stream (an infinite list) of natural numbers:
 - A characterization of the class of Basic Feasible Functionals using term rewrite systems on streams and interpretation methods has been proposed in [13]. This result is part of Hugo Férée's PhD thesis for which he has obtained the Ackermann award.
 - In [14], we have provided some interpretation criteria useful to ensure two kinds of stream properties: space upper bounds and input/output upper bounds. Our space upper bounds criterion ensures global and local upper bounds on the size of each output stream element expressed in term of the maximal size of the input stream elements. The input/output upper bounds criterion considers instead the relations between the number of elements read from the input stream and the number of elements produced on the output stream.
 - The paper [21] has extended the light affine lambda calculus with inductive and coinductive data types using the category theory notions of (weak) initial algebra and coalgebra.
- **Complexity analysis of Object-Oriented programs.** We have proposed a type system based on non-interference and data ramification (tiering) principles in [22]. We have captured the set of functions computable in polynomial time on OO programs. The studied language is general enough to capture most OO constructs and the characterization is quite expressive as it allows the analysis of a combination of imperative loops and of data ramification scheme based on Bellantoni and Cook's safe recursion using function algebra.
- **Rice-like theorem for primitive recursive functions.** We have studied the following question: what are the properties of primitive recursive functions that are decidable (by a Turing machine), given a primitive recursive presentation of the function. We give a complete characterization of these properties. We show that they can be expressed as unions of elementary properties of being compressible. If $h : \mathbb{N} \rightarrow \mathbb{N}$ is a computable increasing unbounded function (like $\log(n)$ or 2^n), we say that a function $f : \mathbb{N} \rightarrow \mathbb{N}$ is h -compressible if for each n there is a primitive recursive program of size at most $h(n)$ computing a function that coincides with f on entries $0, 1, \dots, n$. Whether f is h -compressible is decidable given a primitive recursive program for f , and every decidable property can be obtained as a combination of such elementary properties. This result actually holds for any class of total functions that admits a sound and complete programming language. An article is currently in preparation.
- **Parametrization of geometric figures.** During the master internship of Diego Nava Saucedo, we have studied the semi-computability of geometric figures. A figure is semi-computable if there is a program that semi-decides whether a pixel intersects the figure. Our goal is to understand the semi-computability of a figure in terms of the parameters describing the figure. It turns out that the usual ways of parameterizing simple figures such as triangles, squares or disks do not behave well in terms of semi-computability. We have actually proved that no *finite* parametrization behaves well.
- **Symbolic Dynamics on Groups.** In an effort to better understand the interplay of geometry and computability in tiling theory, E. Jeandel has studied tiling problems on general Cayley graphs, and has obtained a significant number of new results. He has proven that groups with an (strongly) aperiodic tiling system have decidable word problem [30], and provided examples of new groups (in particular monster groups) with such tiling systems, and proved that all nontrivial nilpotent groups

have an aperiodic tiling system and an undecidable domino problem [31]. He also showed how the new concept of translation-like actions from geometric group theory can be used to prove that many groups, in particular the Grigorchuk groups and most groups with a nontrivial center, have an undecidable domino problem [33].

- **The smallest aperiodic tiling set.** In a joint work with Michael Rao, E. Jeandel has proven that there exists an aperiodic set of 11 Wang tiles [34], and furthermore that this number is optimal.

7.2. Quantum Computing

- **On Weak Odd Domination and Graph-based Quantum Secret Sharing.** In this work published in the journal Theoretical Computer Science [15], Simon Perdrix and his co-authors Sylvain Gravier, Jérôme Javelle and Mehdi Mhalla study weak odd domination in graphs and its application in quantum secret sharing. A weak odd dominated (WOD) set in a graph is a subset B of vertices for which there exists a distinct set of vertices C such that every vertex in B has an odd number of neighbors in C . They point out the connections of weak odd domination with odd domination, $[\sigma, \rho]$ -domination, and perfect codes. They introduce bounds on $\kappa(G)$, the maximum size of WOD sets of a graph G , and on $\kappa'(G)$, the minimum size of non-WOD sets of G . Moreover, they prove that the corresponding decision problems are NP-complete. The study of weak odd domination is mainly motivated by the design of graph-based quantum secret sharing protocols: a graph G of order n corresponds to a secret sharing protocol whose threshold is $\kappa_Q(G) = \max(\kappa(G), n - \kappa'(G))$. These graph-based protocols are very promising in terms of physical implementation, however all such graph-based protocols studied in the literature have quasi-unanimity thresholds (i.e. $\kappa_Q(G) = n - o(n)$ where n is the order of the graph G underlying the protocol). In this paper, they show using probabilistic methods the existence of graphs with smaller κ_Q (i.e. $\kappa_Q(G) \leq 0.811n$ where n is the order of G). They also prove that deciding for a given graph G whether $\kappa_Q(G) \leq k$ is NP-complete, which means that one cannot efficiently double check that a graph randomly generated has actually a κ_Q smaller than $0.811n$.
- **Minimum Degree up to Local Complementation: Bounds, Parameterized Complexity, and Exact Algorithms.** In this work presented at ISAAC [25], David Cattaneo and Simon Perdrix introduce new upper bounds and exact algorithms for the local minimum degree. The author also prove the $W[2]$ -membership of the corresponding decision problem. The local minimum degree of a graph is the minimum degree that can be reached by means of local complementation. For any n , there exist graphs of order n which have a local minimum degree at least $0.189n$, or at least $0.110n$ when restricted to bipartite graphs. Regarding the upper bound, they show that the local minimum degree is at most $3/8n + o(n)$ for general graphs and $n/4 + o(n)$ for bipartite graphs, improving the known $n/2$ upper bound. They also prove that the local minimum degree is smaller than half of the vertex cover number (up to a logarithmic term). The local minimum degree problem is NP-Complete and hard to approximate. They show that this problem, even when restricted to bipartite graphs, is in $W[2]$ and FPT-equivalent to the EvenSet problem, whose $W[1]$ -hardness is a long standing open question. Finally, they show that the local minimum degree is computed by a $O_*(1.938n)$ -algorithm, and a $O_*(1.466n)$ -algorithm for the bipartite graphs.
- **The ZX Calculus is incomplete for Clifford+T quantum mechanics.** The ZX calculus is a diagrammatic language for quantum mechanics and quantum information processing. In this paper [17], Simon Perdrix and Hany Wang prove that the ZX-calculus is not complete for the Clifford+T quantum mechanics. The completeness for this fragment has been stated as one of the main current open problems in categorical quantum mechanics. The ZX calculus was known to be incomplete for quantum mechanics, on the other hand, it has been proved complete for Clifford quantum mechanics (a.k.a. stabilizer quantum mechanics), and for single-qubit Clifford+T quantum mechanics. The question of the completeness of the ZX calculus for Clifford+T is a crucial step in the development of the ZX calculus because of its (approximate) universality for quantum mechanics (i.e. any unitary evolution can be approximated using Clifford and T gates only). They exhibit a property which is known to be true in Clifford+T quantum mechanics and prove that this equation cannot be derived

in the ZX calculus, by introducing a new sound interpretation of the ZX calculus in which this particular property does not hold. Finally, we propose to extend the language with a new axiom. This result has been presented as invited speakers in the conferences "Quantum Theory: from foundations to technologies" in Vaxjo Sweden, and "Higher TQFT and categorical quantum mechanics" at the Scrounger Institute in Vienna. The authors also presented these results at the workshop of the CNRS groupe de travail Informatique Quantique du GDR IM, in Grenoble.

- **Block Representation of Reversible Causal Graph Dynamics.** In this work presented at the conference on Foundation of computer science (FCT'15) [18], Pablo Arrighi, Simon Martiel and Simon Perdrix, consider a reversible version of the causal graph dynamics. Causal Graph Dynamics extend Cellular Automata to arbitrary, bounded-degree, time-varying graphs. The whole graph evolves in discrete time steps, and this global evolution is required to have a number of physics-like symmetries: shift-invariance (it acts everywhere the same) and causality (information has a bounded speed of propagation). We study a further physics-like symmetry, namely reversibility. More precisely, we show that Reversible Causal Graph Dynamics can be represented as finite-depth circuits of local reversible gates.
- **Reversibility in the Extended Measurement-based Quantum Computation.** In this work by Nidal Hamrit and Simon Perdrix has been presented at the conference on Reversible Computation in Grenoble [23]. When applied on some particular quantum entangled states, measurements are universal for quantum computing. In particular, despite the fundamental probabilistic evolution of quantum measurements, any unitary evolution can be simulated by a measurement-based quantum computer (MBQC). They consider the extended version of the MBQC where each measurement can occur not only in the X,Y-plane of the Bloch sphere but also in the X,Z- and Y,Z-planes. The existence of a gflow in the underlying graph of the computation is a necessary and sufficient condition for a certain kind of determinism. They extend the focused gflow (a gflow in a particular normal form) defined for the X,Y-plane to the extended case, and provide necessary and sufficient conditions for the existence of such normal forms.
- **Quantum Circuits for the Unitary Permutation Problem.** In this paper presented at TAMC'15 [20] Stefano Facchini and Simon Perdrix consider the *Unitary Permutation* problem which consists, given n quantum gates U_1, \dots, U_n and a permutation σ of $\{1, \dots, n\}$, in applying the quantum gates in the order specified by σ , i.e., in performing $U_{\sigma(n)} \circ \dots \circ U_{\sigma(1)}$. This problem has been introduced and investigated in [47] where two models of computations are considered. The first is the (standard) model of query complexity: the complexity measure is the number of calls to any of the quantum gates U_i in a quantum circuit which solves the problem. The second model is roughly speaking a model for higher order quantum computation, where quantum gates can be treated as objects of second order. In both model the existing bounds are improved, in particular the upper and lower bounds for the standard quantum circuit model are established by pointing out connections with the *permutation as substring* problem introduced by Karp.

CASCADE Project-Team

6. New Results

6.1. Results

All the results of the team have been published in journals or conferences (see the list of publications). They are all related with the research program (see before) and the research projects (see after):

- New zero-knowledge proofs
- Advanced families of hash proofs
- More efficient constructions with lattices
- New e-cash constructions
- Advanced primitives for the privacy in the cloud
- Efficient functional encryption
- Various predicate encryption schemes
- Cryptanalysis of symmetric primitives
- New leakage-resilient primitives
- Stronger security with related-key security

CASSIS Project-Team

7. New Results

7.1. Automated Deduction

We develop general techniques which allow us to re-use available tools in order to build a new generation of solvers offering a good trade-off between expressiveness, flexibility, and scalability. We focus on the careful integration of combination techniques and rewriting techniques to design decision procedures for a wide range of verification problems.

7.1.1. *Building and Verifying decision procedures*

Participants: Alain Giorgetti, Olga Kouchnarenko, Christophe Ringeissen.

In the context of the PhD thesis by Elena Tushkanova (defended in 2013), we have developed a methodology to build decision procedures specified by using a superposition calculus [20] which is at the core of all equational theorem provers. This calculus is refutation complete: it provides a semi-decision procedure that halts on unsatisfiable inputs but may diverge on satisfiable ones. Fortunately, it may also terminate for some theories of interest in verification, and thus it becomes a decision procedure. To reason on the superposition calculus, a schematic superposition calculus has been developed to build the schematic form of the saturations allowing to automatically prove decidability of single theories and of their combinations. We have proposed a rule-based logical framework and a tool implementing a complete many-sorted schematic superposition calculus for arbitrary theories. By providing results for unit theories, arbitrary theories, and also for theories with counting operators, we show that this tool is very useful to derive decidability and combinability of theories of practical interest in verification.

7.1.2. *Combination of Satisfiability Procedures*

Participant: Christophe Ringeissen.

We have continued our work started with Paula Chocron (IIIA-CSIC, U. Barcelona) and Pascal Fontaine (project-team Veridis) on the extension of the Nelson-Oppen combination method to non-disjoint unions of theories. We investigate the case of theories connected via bridging functions [28]. The motivation is, e.g., to solve verification problems expressed in a combination of data structures connected to arithmetic with bridging functions such as the length of lists and the size of trees. We present a sound and complete combination procedure à la Nelson-Oppen for the theory of absolutely free data structures, including lists and trees. This combination procedure is then refined for standard interpretations. The resulting theory has a nice politeness property, enabling combinations with arbitrary decidable theories of elements.

To go beyond the case of absolutely free data structures, we study in [29] the problem of determining the data structure theories for which this combination method is sound and complete. Our completeness proof is based on a rewriting approach where the bridging function is defined as a term rewrite system, and the data structure theory is given by a basic congruence relation. Our contribution is to introduce a class of data structure theories that are combinable with a disjoint target theory via an inductively defined bridging function. This class includes the theory of equality, the theory of absolutely free data structures, and all the theories in between. Hence, our non-disjoint combination method applies to many classical data structure theories admitting a rewrite-based satisfiability procedure.

7.1.3. *Unification Modulo Equational Theories*

Participant: Christophe Ringeissen.

We investigate a hierarchical combination approach to the unification problem in non-disjoint unions of equational theories. In this approach, the idea is to extend a base theory with some additional axioms given by rewrite rules in such way that the unification algorithm known for the base theory can be reused without loss of completeness. Additional techniques are required to solve a combined problem by reducing it to a problem in the base theory. In [33] we show that the hierarchical combination approach applies successfully to some classes of syntactic theories, such as shallow theories since the required unification algorithms needed for the combination algorithm can always be obtained. We also consider the matching problem in syntactic extensions of a base theory. Due to the more restricted nature of the matching problem, we obtain several improvements over the unification problem.

7.1.4. Enumeration of Planar Proof Terms

Participant: Alain Giorgetti.

By the Curry-Howard isomorphism, simply typed lambda terms correspond to natural deduction proofs in minimal logic. Noam Zeilberger and Alain Giorgetti proved that normal planar lambda terms are in size-preserving bijection with rooted planar maps [21]. Although the formal aspect is not emphasized in the paper, the use of formal representations of both normal planar lambda terms and rooted planar maps, of logic programming and a proof assistant software helped much in more quickly finding the bijection.

7.1.5. Rewriting-based Mathematical Model Transformations

Participants: Walid Belkhir, Alain Giorgetti.

Since 2011 we collaborate with the Department “Temps-Fréquence” of the FEMTO-ST institute (Franche-Comté Electronique Mécanique Thermique et Optique - Sciences et Technologies, CNRS UMR 6174) on the formalization of asymptotic methods (based on two-scale convergence). The goal is to design a software, called *MEMSALab*, for the automatic derivation of multiscale models of arrays of micro- and nanosystems. In this domain a model is a partial differential equation. Multiscale methods approximate it by another partial differential equation which can be numerically simulated in a reasonable time. The challenge consists in taking into account a wide range of different physical features and geometries e.g. thin structures, periodic structures, multiple nested scales etc. In [24], we propose a method called “*by-extension-combination*”, in which the asymptotic models are constructed incrementally so that model features can be included step by step. More precisely, a model derivation is formalised as a rewriting strategy, and its extension is formalised as a second-order rewriting strategy. Thus, our method amounts to defining an operation of combination over a class of second-order rewriting strategies. We illustrate the method by an example of an asymptotic model for the stationary heat equation in a Micro-Mirror Array developed for astrophysics.

7.2. Security Protocol Verification

The design of cryptographic protocols is error-prone. Without a careful analysis, subtle flaws may be discovered several years after the publication of a protocol, yielding potential harmful attacks. In this context, formal methods have proved their interest for obtaining good security guarantees. Many analysis techniques have been proposed in the literature [66]. We have edited a book [62] where each chapter presents an important and now standard analysis technique. We develop new techniques for richer primitives, wider classes of protocols and higher security guarantees. In Section 7.4.3 we consider derived testing techniques for verifying protocol implementations.

7.2.1. Design of Voting Protocols

Participants: Véronique Cortier, Stéphane Glondu, Steve Kremer, Peter Rønne.

Voting is a cornerstone of democracy and many voting systems have been proposed so far, from old paper ballot systems to purely electronic voting schemes. Although many works have been dedicated to standard protocols, very few address the challenging class of voting protocols.

One famous e-voting protocol is Helios, an open-source web-based end-to-end verifiable electronic voting system, used e.g., by UCL and the IACR association in real elections. One main advantage of Helios is its verifiability, up-to the ballot box (a dishonest ballot box may add ballots). We have defined a variant of Helios, named Belenios, that prevents from ballot stuffing, even against a dishonest ballot box. Our approach consists in introducing an additional authority that provides credentials that the ballot box can verify but not forge. Belenios⁰ has been implemented by Stéphane Glondu (cf Section 6.1.3).

Helios as well as Belenios are not receipt-free, that is, a (malicious) voter can *prove* how they voted to any third party. Building upon a scheme proposed by G. Fuschbauer and David Pointcheval, we have enhanced Belenios with a receipt-free variant, called BeleniosRF. Now, the ballot box can re-randomize any (signed) ballot it receives. This way, a voter can no longer exhibit the randomness they used to build their ballot.

End-to-end verifiable voting schemes typically involves voters handling an encrypted ballot in order to confirm that their vote is accurately included in the tally. While this may be technically valid, from a public acceptance standpoint it may be problematic: many voters may not really understand the purpose of the encrypted ballot and the various checks that they can perform. In [61] we take a different approach and revisit an old idea: to provide each voter with a private tracking number. Votes are posted on a bulletin board in the clear along with their associated tracking number. This is appealing in that it provides voters with a very simple, intuitive way to verify their vote, in the clear. However, there are obvious drawbacks: we must ensure that no two voters are assigned the same tracker and we need to keep the trackers private. We propose a new scheme, called Selene, that addresses both of these problems: we ensure that voters get unique trackers and we close off the coercer's window of opportunity by ensuring that the voters only learn their tracking numbers after votes have been posted. The resulting scheme provides receipt-freeness, and indeed a good level of coercion-resistance while also providing a more immediately understandable form of verifiability. The cryptography is under the bonnet as far as the voter is concerned.

In 2010 Hao, Ryan and Zielinski proposed a simple decentralised e-voting protocol that only requires 2 rounds of communication. Thus, for k elections their protocol needs $2k$ rounds of communication. Observing that the first round of their protocol is aimed to establish the public-keys of the voters, we propose in [60] an extension of the protocol as a non-interactive e-voting scheme in the public-key setting (NIVS) in which the voters, after having published their public-keys, can use the corresponding secret-keys to participate in an arbitrary number of one-round elections. We first construct a NIVS with a standard tally function where the number of votes for each candidate is counted. Further, we present constructions for two alternative types of elections. Specifically in the first type (dead or alive elections) the tally shows if at least one voter cast a vote for the candidate. In the second one (elections by unanimity), the tally shows if all voters cast a vote for the candidate. Our constructions are based on bilinear groups of prime order. As definitional contribution we provide formal computational definitions for privacy and verifiability of NIVSs. We conclude by showing intriguing relations between our results, secure computation, electronic exams and conference management systems.

7.2.2. Analysis of Voting Protocols

Participants: Véronique Cortier, Catalin Dragan, Steve Kremer, Peter Rønne.

Properties. Even a basic property like ballot secrecy is difficult to define formally and several definitions co-exist. We studied all game-based privacy definitions of the literature and discovered that none of them was satisfactory: they were either limited (not fully modeling e-voting protocols), or too strong (incompatible with verifiability), or even flawed for a few of them [25]. Based on our findings, we have proposed a new game-based privacy definition BPRIV, proved that it implies simulation-based privacy and showed that it is realized by the Helios protocol [25].

Proof. Such a proof of privacy for Helios is done by hand and is error-prone. Moreover, there is not a single version of Helios. Instead, many slight variants of Helios may be considered (e.g. early and late weeding, weeding based on the identity or on the ciphertexts, mixnet or homomorphic tally, etc.). Each of these variants would require a new proof. Therefore, we are conducting a proof of Helios and Belenios through the Easycrypt framework. This first fully formal proof will cover most existing variants of Helios and Belenios.

⁰<https://belenios.loria.fr>

Analysis. Existing automated analysis techniques are inadequate to deal with commonly used cryptographic primitives, such as homomorphic encryption and mix-nets, as well as some fundamental security properties, such as verifiability. In collaboration with Matteo Maffei and Fabienne Eigner (Saarland University) we propose a novel approach based on refinement type systems for the automated analysis of two fundamental properties of e-voting protocols, namely, vote privacy and verifiability. We demonstrate the effectiveness of our approach by developing the first automated analysis of Helios using an off-the-shelf type-checker [32].

A challenging problem in e-voting is to provide guarantees when the voting platform itself is corrupted. Du-Vote [73] is a recently presented remote electronic voting scheme that aims to be malware tolerant, i.e., provide security even in the case where the platform used for voting has been compromised by dedicated malware. For this it uses an additional hardware token, similar to tokens distributed in the context of online banking. Du-Vote aims at providing vote privacy as long as either the vote platform or the vote server is honest. For verifiability, the security guarantees are even higher, as even if the token's software has been changed, and the platform and the server are colluding, attempts to change the election outcome should be detected with high probability. In recent work [41] we provide an extensive security analysis of Du-Vote and show several attacks on both privacy as well as verifiability. We also propose changes to the system that would avoid many of these attacks.

7.2.3. Other Families of Protocols

Participants: Véronique Cortier, Jannik Dreier, Alicia Filipiak, Steve Kremer, Ludovic Robin.

Secure Mobile Applications. There is a growing development of Secure Elements for Mobile Phone and Tablets. These Secure Elements are hosted in the SIM for example and can perform cryptographic operations. This opens the way for a much higher level of security in such environments. However, how to use these secure elements is still very unclear. How keys will be registered in Secure Elements? Which applications may access to the keys and how is this enforced? Which part of the application should be deployed in a Secure Element? It is of course not possible to host an entire application in a Secure Element for size and performance issues. Alicia Filipiak has started a PhD in March 2015 to propose a model for secure mobile applications that make use of Secure Elements. This is a collaboration with Orange Labs (CIFRE). She has proposed a light and secure paiement application which is compatible with standard paiement systems (EMV). The proof of security is conducted in Tamarin, in order to cope with global states.

Protocols using low-entropy secrets. Many two factor authentication protocols consider an additional authentic, but low bandwidth channel to send a confirmation code. A typical example is to send such a code by SMS to a user's mobile phone. Given that such codes need to be copied by users they are short and therefore vulnerable to offline brute-force attacks. Ludovic Robin has started a PhD thesis in October 2014 and proposed a model to take into account an attacker's capability to run such brute-force attacks. While the problem is reminiscent to guessing attacks in password-based protocols, several subtle differences make this problem more difficult. Ludovic is adapting the decision procedure implemented in *Akiss* in order to decide protocol security in the presence of such an attacker.

Auction protocols. Auctions have a long history, having been recorded as early as 500 B.C.. Nowadays, electronic auctions have been a great success and are increasingly used in various applications. Many cryptographic protocols have been proposed to address the various security requirements of these electronic transactions, in particular to ensure privacy. Jannik Dreier, in collaboration with Pascal Lafourcade from Université d'Auvergne and Jean-Guillaume Dumas from Université Grenoble Alpes, recently performed a detailed analysis [15] of Brandt's auction protocol that computes the winner using homomorphic operations on a distributed ElGamal encryption of the bids. Jannik and his coauthors were able to show that this protocol – when using malleable interactive zero-knowledge proofs – is vulnerable to attacks by dishonest bidders. Such bidders can manipulate the publicly available data in a way that allows the seller to deduce all participants' bids. They developed an efficient parallelized implementation of the protocol and the attack to show its practicality.

7.2.4. Automated Verification of Indistinguishability Properties

Participants: Vincent Cheval, Rémy Chréten, Véronique Cortier, Antoine Dallon, Jannik Dreier, Steve Kremer.

New emerging classes of protocols such as voting protocols often require to model less classical security properties, such as anonymity properties, strong versions of confidentiality and resistance to offline guessing attacks. Many of these properties can be modelled using the notion of indistinguishability by an adversary, which can be conveniently modeled using process equivalences.

Active case, bounded number of sessions. We previously proposed a procedure for approximating trace equivalence in the case of a bounded number of sessions, i.e., for a replication free fragment of a cryptographic process calculus. The procedure is implemented in the *Akiss* tool. While we proved soundness and correctness for any convergent rewrite system that has the finite variant property, termination of the procedure was still an open question. We have recently shown that the procedure indeed terminates for the class of subterm convergent rewrite systems. We are currently also working on an extension of *Akiss* in order to verify protocols that may use the exclusive or operator. This extension requires us to reason modulo associativity and commutativity. While proving soundness and completeness of a naive extension of the existing procedure is a rather straightforward, the resulting procedure faces directly non-termination. We therefore adapt the resolution strategy to ensure termination on practical examples. While soundness is preserved we need to prove the completeness of the new resolution strategy.

When considering the equational theory corresponding to the standard primitives, Vincent Cheval has proposed a decision procedure for checking equivalence of set constraints, which yields a procedure for checking trace equivalence [69]. We have extended this decision procedure to the case where the attacker can observe the *time* of executions [27], capturing what is called *timing attacks*. To obtain decidability, we have shown how to reduce to a previous result to decide length trace equivalence, where the attacker no longer has access to execution times but can still compare the length of messages. As an application, we study several protocols that aim for privacy. In particular, we (automatically) detect an existing timing attack against the biometric passport and new timing attacks against the Private Authentication protocol.

Active case, unbounded number of sessions.

We have shown that for some classes of protocols, decidability of trace equivalence can be reduced to equivalence of deterministic pushdown automata [13]. Equivalence of deterministic pushdown automata is decidable [79] and the corresponding decision procedure has been recently implemented by Géraud Senizergues. Based on his tool, we have developed a tool for automatically checking equivalence, for an unbounded number of sessions.

For trace properties such as secrecy and authentication, it has been shown that it is sufficient to consider typically three agents, two honest and one dishonest agents [70]. This result no longer holds for equivalence properties. Antoine Dallon has recently started a PhD thesis on deciding equivalence properties. He has shown that it is sufficient to consider two honest agents and two dishonest agents for equivalence properties, for deterministic processes with standard primitives and without else branches. More generally, he shows how to bound the number of agents for arbitrary constructor theories and for protocols with simple else branches. These hypotheses are tight, and counter-examples are provided for non action-deterministic processes, non constructor theories, or protocols with complex else branches.

When proving security in symbolic settings for an unbounded number of sessions, a typical technique (like in the aforementioned result) consists in abstracting away fresh nonces and keys by a bounded set of constants. While this abstraction is clearly sound in the context of secrecy properties (for protocols without else branches), this is no longer the case for equivalence properties. We have shown how to soundly get rid of nonces in the context of equivalence properties [30]. We show that nonces can be replaced by constants provided that each nonce is associated to two constants (instead of typically one constant for secrecy properties). Our result holds for deterministic (simple) protocols and a large class of primitives that includes e.g. standard primitives, blind signatures, and zero-knowledge proofs.

Of course, our abstraction of nonces may introduce false attacks. To avoid this, it is necessary to consider protocols *with* nonce. We have provide the first decidability result for trace equivalence of security protocols, for an unbounded number of sessions and unlimited fresh nonces [31]. Our class encompasses most symmetric key protocols of the literature, in their tagged variant.

Decomposing equivalence. Unique decomposition has been a subject of interest in process algebra for a long time (for example in BPP or CCS in the 1980s), as it provides a normal form and useful cancellation properties. In recent work [16] Jannik Dreier, together with Cristian Ene and Yassine Lakhnech from Université Grenoble Alpes as well as Pascal Lafourcade from Université d'Auvergne, proved two parallel decomposition results for subsets of the applied π -calculus. They showed that every closed normed (i.e. with a finite shortest complete trace) process P can be decomposed uniquely into prime factors P_i with respect to strong labeled bisimilarity, i.e. such that $P \sim_l P_1 | \dots | P_n$. Moreover, they proved that closed finite processes can be decomposed uniquely with respect to weak labeled bisimilarity. They also investigated whether efficient algorithms that compute the unique decompositions exist, which would be useful for the verification of equivalences. It turned out that the simpler problem of deciding whether a process is in its unique decomposition form is undecidable in general in both cases, due to potentially undecidable equational theories. Moreover, the unique decomposition remains undecidable even given an equational theory with a decidable word problem.

7.2.5. Securely Composing Protocols

Participants: Vincent Cheval, Véronique Cortier, Éric Le Morvan.

Protocols are often built in a modular way. For example, authentication protocols may assume pre-distributed keys or may assume secure channels. However, when an authentication protocol has been proved secure assuming pre-distributed keys, there is absolutely no guarantee that it remains secure when executing a real protocol for distributing the keys. During his PhD thesis, Éric Le Morvan has shown how to securely realize the three main types of channels: secure (unreadable and untappable), confidential (unreadable), and authenticated (untappable) channels [54].

7.3. Model-based Verification

We have investigated extensions of regular model-checking to new classes of rewrite relations on terms. We have studied specification and proof of modular imperative programs, as well as of modal workflows.

7.3.1. Tree Automata with Constraints

Participants: Pierre-Cyrille Héam, Olga Kouchnarenko.

Tree automata with constraints are widely used to tackle data base algorithmic problems, particularly to analyse queries over XML documents. The model of Tree Automata with Global Constraints (TAGED) has been introduced for these purposes. The membership problem for TAGED is known to be NP-complete. The emptiness problem for TAGED is known to be decidable and the best known algorithm in the general case is non elementary. Following our NP-hardness result [74], we are still working in collaboration with Vincent Hugot on the complexity of the emptiness problem.

7.3.2. Random Generation of Finite Automata

Participant: Pierre-Cyrille Héam.

Developing new algorithms and heuristics raises crucial evaluation issues, as improved worst-case complexity upper-bounds do not always transcribe into clear practical gains. A classical way for software performance evaluation is to randomly generate inputs.

In collaboration with Jean-Luc Joly, we investigate the problem of randomly and uniformly generating deterministic pushdown automata [40]. Based on a recursive counting approach, we propose a polynomial time algorithm for this purpose. The influence of the accepting condition on the generated automata is also experimentally studied.

Partially ordered automata are finite automata where simple loops have length one. We have used a Markov chain based approach [75] to randomly - and uniformly - generate deterministic partially ordered automata.

In [39] we address the problem of the uniform random generation of non deterministic automata (NFA) up to isomorphism. We show how to use a Monte-Carlo approach to uniformly sample a NFA. The main result is to show how to use the Metropolis-Hastings Algorithm to uniformly generate NFAs up to isomorphism. Using labeling techniques, we show that in practice it is possible to move into the modified Markov Chain efficiently, allowing the random generation of NFAs up to isomorphism with dozens of states. This general approach is also applied to several interesting subclasses of NFAs (up to isomorphism), such as NFAs having a unique initial states and a bounded output degree. Finally, we prove that for these interesting subclasses of NFAs, moving into the Metropolis Markov chain can be done in polynomial time.

7.3.3. Verification of Linear Temporal Patterns over Finite and Infinite Traces

Participants: Pierre-Cyrille Héam, Olga Kouchnarenko.

In the regular model-checking framework, reachability analysis can be guided by temporal logic properties, for instance to achieve the counter example guided abstraction refinement (CEGAR) objectives. A way to perform this analysis is to translate a temporal logic formula expressed on maximal rewriting words into a “rewrite proposition” – a propositional formula whose atoms are language comparisons, and then to generate semi-decision procedures based on (approximations of) the rewrite proposition. In collaboration with Vincent Hugot, we have investigated suitable semantics for LTL on maximal rewriting words and their influence on the feasibility of a translation. We have expended the work in [76] by providing a general translation scheme giving exact results for a fragment of LTL corresponding mainly to safety formulæ, and approximations for a larger fragment.

7.3.4. Constraint Solving for Verifying Modal Workflow Specifications

Participants: Hadrien Bride, Olga Kouchnarenko.

Workflow Petri nets are well suited for modelling and analysing discrete event systems exhibiting behaviours such as concurrency, conflict, and causal dependency between events. They represent finite or infinite-state processes, and several important verification problems, like reachability or soundness, are known to be decidable. Modal specifications introduced in [77] allow loose or partial specifications in a framework based on process algebras.

Our work in [26] aims at verifying modal specifications of coloured workflows with data assigned to the tokens and modified by transitions. To this end, executions of coloured workflow nets are modelled using constraint systems, and constraint solving is used to verify modal specifications specifying necessary or admissible behaviours. An implementation supporting the proposed approach and promising experimental results on an issue tracking system constitute a practical contribution.

7.4. Model-based Testing

Our research in Model-Based Testing (MBT) aims to extend the coverage of tests. The coverage refers to several artefacts: model, test scenario/property, and code of the program under test. The test generation uses various underlying techniques such as symbolic animation of models [71], or symbolic execution of programs by means of dedicated constraints, SMT solvers, or model-checkers.

7.4.1. Automated Test Generation from Behavioral Models

Participants: Fabrice Bouquet, Frédéric Dadeau, Elizabetha Fourneret, Jean-Marie Gauthier, Julien Lorrain, Alexandre Vernotte.

We have developed an original model-based testing approach that takes a behavioral view (modelled in UML) of the system under test and automatically generates test cases and executable test scripts according to model coverage criteria. We continue to extend this result to SysML specifications for validating embedded systems [35]. We apply this method on smartSurface [34]

We have investigated the use of a model-based testing approach for vulnerability testing in web applications. Our research prototype was able to detect vulnerabilities on already deployed web applications [80].

7.4.2. Scenario-Based Verification and Validation

Participants: Fabrice Bouquet, Frédéric Dadeau, Elizabeta Fourneret.

Test scenarios represent an abstract test case specification that aims at guiding the model animation in order to produce relevant test cases. Contrary to the previous section, this technique is not fully automated since it requires the user to design the scenario, in addition to the model.

We have proposed a dedicated formalism to express test properties. A test property is first translated into a finite state automaton which describes a monitor of its behaviors. We have also proposed dedicated property coverage criteria that can be used either to measure the property coverage of a given test suite, or to generate test cases, exercising nominal or robustness aspects of the property. This process has been fully tool-supported into an integrated software prototype⁰. This process has been designed during the ANR TASCOC project (2009-2012) and was continued during the ANR ASTRID OSEP project (2012-2013). The industrialization of this approach, and its integration within commercial test generation tools has started with the ANR ASTRID Maturation MBT_Sec project (2014-2015) in collaboration with the French DoD [46]. A technology transfer is currently in progress to integrate this technology into the Smartesting CertifyIt test generation environment.

Also, we have experimented the model approach to validate and to design Multi-Agent systems [51], [52].

7.4.3. Mutation-based Testing of Security Protocols

Participants: Frédéric Dadeau, Pierre-Cyrille Héam, Michaël Rusinowitch.

We have proposed a model-based penetration testing approach for security protocols [14]. This technique relies on the use of mutations of an original protocol, proved to be correct, for injecting realistic errors that may occur during the protocol implementation (e.g., re-use of existing keys, partial checking of received messages, incorrect formatting of sent messages, use of exponential/xor encryption, etc.). Mutations that lead to security flaws are used to build test cases, which are defined as a sequence of messages representing the behavior of the intruder. We have applied our technique on protocols designed in HLPSL, and implemented the protocol mutation tool jMuHLPSL that performs the mutations. The mutants are then analyzed by *CL-AtSe*.

7.4.4. Code and Contract-based Test Generation and Static Analysis

Participants: Fabrice Bouquet, Frédéric Dadeau, Alain Giorgetti.

With the CEA we have developed a test generation technique based on C code and formal specifications, to facilitate deductive verification, in a new tool named StaDy [43]. The tool integrates the concolic test generator PathCrawler within the static analysis platform Frama-C. StaDy is able to handle the ANSI C Specification Language (ACSL) of the framework and other Frama-C plug-ins are able to reuse results from the test generator. This tool is designed to be the foundation stone of modular static and dynamic analysis combinations in the Frama-C platform.

For bounded exhaustive unitary testing of functions on structured arrays we have designed and formally verified with Frama-C a library of sequential generators [43], [36]. A structured array is an array satisfying given constraints, such as being sorted or having no duplicate values. A sequential generator of structured arrays can be defined by two C functions: the first one computes an initial array, and the second one steps from one array to the next one according to some total order on the set of arrays. We formally specify with ACSL annotations that the generated arrays satisfy the prescribed structural constraints (soundness property) and that the generation is in increasing lexicographic order (progress property). We refine this specification into two programming and specification patterns: one for generation in lexicographic order and one for generation by filtering the output of another generator. After adding suitable loop invariants we automatically prove the soundness and progress properties of many generators with the Frama-C platform.

⁰A video of the prototype is available at: <http://vimeo.com/53210102>

7.5. Verification of Collaborative Systems

We investigate security problems occurring in decentralized systems. We develop general techniques to enforce read and update policies for controlling access to XML documents based on recursive DTDs (Document Type Definition). Moreover, we provide a necessary and sufficient condition for undoing safely replicated objects in order to enforce access control policies in an optimistic way. We investigate privacy issues for social networks in order to give more control to users over their personal data.

7.5.1. Automatic Analysis of Web Services Security

Participants: Walid Belkhir, Michaël Rusinowitch, Mathieu Turuani, Laurent Vigneron.

Automatic composition of web services is a challenging task. Many works have considered simplified automata models that abstract away from the structure of messages exchanged by the services. For the domain of secured services (using e.g., digital signing or timestamping) we have proposed an original approach to automated orchestration of services under security constraints. Given a community of services and a goal service, we reduce the problem of generating a mediator between a client and a service community to a security problem where an intruder should intercept and redirect messages from the service community and a client service till reaching a satisfying state.

In [12] we develop an alternative approach based on *parametrized automata*, a natural extension of finite-state automata over infinite alphabet. In this model the transitions are labeled with constants or variables that can be refreshed in some specified states. We show the applicability of our model to Web services handling data from an infinite domain. We reduce the Web service composition problem to the construction of a simulation of a target service by the asynchronous product of existing services, and prove that this construction is computable. We also show expressive equivalence and succinctness of parametrized automata with respect to Finite Memory Automata in [47] We now work on synthesizing composed services that satisfy required security policies.

7.5.2. Querying Security Views over XML Data

Participant: Abdessamad Imine.

To enforce access control over XML data, virtual security views are commonly used in many many applications and commercial database systems. Querying these views raises some serious problems. More precisely, user XPath queries posed on recursive views cannot be rewritten to be evaluated on the underlying XML data. Existing rewriting solutions are based on the non-standard language “Regular XPath” enabling recursion operator. However, query rewriting under Regular XPath can be of exponential size. In [17], we show that query rewriting is always possible for arbitrary security views (recursive or not) by using only the expressive power of the standard XPath. We propose a more expressive language to specify XML access control policies as well as an efficient algorithm to enforce such policies. Finally, we present our system, called SVMAX, that implements our solutions and we show that it scales well through an extensive experimental study based on real-life DTD.

7.5.3. Secure Computation in Social Networks

Participants: Younes Abid, Bao Thien Hoang, Abdessamad Imine, Huu Hiep Nguyen, Michaël Rusinowitch.

Online social networks are increasingly exploited as real platforms for creating social links and sharing data. They are used from organizing public opinion polls about any societal theme to publish social graph data for achieving in-depth studies. To securely perform these large-scale computations, we need the design of reliable protocols to ensure the data privacy. In [44], [9], we address the polling problem in social networks where users want to preserve the confidentiality of their votes, obtain the correct final result, and hide, if any, their misbehaviors. We present EPol, a simple decentralized polling protocol that is deployed on a family of social graphs that satisfy a property based on topological ordering. Using these graphs, we show that their structures enable low communication cost, ensure vote privacy and limit the impact of dishonest users on the accuracy of the polling output.

The problem of private publication of social graphs has attracted a lot of attention recently. In [50], we tackle the problem about the upper bounds of privacy budgets related to differentially private release of graphs. We provide such a bound and we prove that, with a privacy budget of $O(\log n)$, there exists an algorithm capable of releasing a noisy output graph with edge edit distance of $O(1)$ against the true graph. At the same time, the complexity of our algorithm *Top - m Filter* is linear in the number of edges m . This lifts the limits of the state-of-the-art, which incur a complexity of $O(n^2)$ where n is the number of nodes and runnable only on small graphs.

Anonymous use of Social network do not prevent users from privacy risks resulting from inferring and cross-checking information published by themselves or their relationships. In [57], we have conducted a survey in order to measure sensitiveness of personal data published on social media and to analyze the users behaviors. We have shown that 76% of internet users that have answered the survey are vulnerable to identity or sensitive data disclosure.

7.5.4. *Safe Protocols for Collaborative Applications*

Participant: Abdessamad Imine.

The Operational Transformation (OT) approach, used in many collaborative editors, allows a group of users to concurrently update replicas of a shared object and exchange their updates in any order. The basic idea is to transform any received update operation before its execution on a replica of the object. Designing transformation functions for achieving convergence of object replicas is a critical and challenging issue. In this work, we investigate the existence of transformation functions [19]. From the theoretical point of view, two properties, named TP1 and TP2, are necessary and sufficient to ensure convergence. Using controller synthesis technique, we show that there are some transformation functions, which satisfy TP1 for the basic signatures of insert and delete operations. But, there is no transformation function, which satisfies both TP1 and TP2. Consequently, a transformation function which satisfies both TP1 and TP2 must necessarily have additional parameters in the signatures of some update operations. Accordingly, we provide a new transformation function and show formally that it ensures convergence.

CELTIQUE Project-Team

6. New Results

6.1. Certified compilation

We thrive at improving the technology of certified compilation. Our work builds on the infrastructure provided by the CompCert compiler. We are working both at improving the guarantees provided by certified compilation and at formalising state-of-the-art optimisation techniques.

6.1.1. Safer CompCert

Participants: Sandrine Blazy, Frédéric Besson, Pierre Wilke.

The CompCert compiler is proved with respect to an abstract semantics. In previous work [52], we propose a more concrete memory model for the CompCert compiler [68]. This model gives a semantics to more programs and lift the assumption about an infinite memory. This model makes CompCert safer because more programs are captured by the soundness theorem of CompCert and because it allows to reason about memory consumption.

We are investigating the consequences this model on different compiler passes of CompCert [32]. As a sanity check, we prove formally that the existing memory model is an abstraction of our more concrete model thus validating formally the soundness of CompCert's abstract semantics of pointers. We have also port the front-end of the compiler to our new semantics and are working on the compiler back-end.

6.1.2. Verification of optimization techniques

Participants: Sandrine Blazy, Delphine Demange, Yon Fernandez de Retana, David Pichardie.

The CompCert compiler foregoes using SSA, an intermediate representation employed by many compilers that enables writing simpler, faster optimizers. In previous work [51], we have proposed a formally verified SSA-based middle-end for CompCert, addressing two problems raised by Leroy in 2009: giving an intuitive formal semantics to SSA, and leveraging its global properties to reason locally about program optimizations. Since then, we have studied in more depth the SSA-based optimization techniques with the objective to make the middle-end more realistic, in terms of the efficiency of optimizations, and to rationalize the way the correctness proofs of optimizations are conducted and structured.

First, we have studied in [34] the problem of a verified, yet efficient (i.e. as implemented in production compilers) technique for testing the dominance relation between two nodes in a control flow graph. We propose a formally verified validator of untrusted dominator trees, on top of which we implement and prove correct a fast dominance test.

Second, in [20], we implement and verify two prevailing SSA optimizations (Sparse Conditional Constant Propagation and Global Value Numbering), conducting the proofs in a unique and common sparse optimization proof framework, factoring out many of the dominance-based reasoning steps required in proofs of SSA-based optimizations. Our experimental evaluations indicate both a better precision, and a significant compilation time speedup.

Finally, we have studied (paper under review at the international conference Compiler Construction 2016) the destruction of the SSA form (i.e. at the exit point of the middle-end), a problem that has remained a difficult problem, even in a non-verified environment. We formally defined and proved the properties of the generation of Conventional SSA (CSSA) which make its destruction simple to implement and prove. We implemented and proved correct a coalescing destruction of CSSA, à la Boissinot et al., where variables can be coalesced according to a refined notion of interference. Our CSSA-based, coalescing destruction allows us to coalesce more than 99% of introduced copies, on average, and leads to encouraging results concerning spilling and reloading during post-SSA allocation.

6.2. Certified Static Analyses

6.2.1. *Certified Analyses for JavaScript*

Participants: Martin Bodin, Thomas Jensen, Alan Schmitt.

We have continued our work on the certification of analyses for JavaScript by developing a systematic way to build certified abstract interpreters from big-step semantics using the Coq proof assistant. We based our approach on Schmidt's abstract interpretation principles for natural semantics, and used a pretty-big-step (PBS) semantics, a semantic format proposed by Charguéraud. We proposed a systematic representation of the PBS format and implemented it in Coq. We then showed how the semantic rules can be abstracted in a methodical fashion, independently of the chosen abstract domain, to produce a set of abstract inference rules that specify an abstract interpreter. We proved the correctness of the abstract interpreter in Coq once and for all, under the assumption that abstract operations faithfully respect the concrete ones. We finally showed how to define correct-by-construction analyses: their correction amounts to proving they belong to the abstract semantics. This work has been published at CPP 2015 [19].

In addition, we have worked on hybrid typing of information flow for JavaScript, in collaboration with José Fragoso Santos and Tamara Rezk at Inria Sophia-Antipolis. Our analysis combined static and dynamic typing in order to avoid rejecting programs due to imprecise typing information. This work has been published at TGC 2015 [21].

6.2.2. *Certified Analyses for safety-critical C programs*

Participants: Sandrine Blazy, Vincent Laporte, David Pichardie.

We designed and proved sound, using the Coq proof assistant, a static analyzer, Verasco [26], based on abstract interpretation for most of the ISO C 1999 language (excluding recursion and dynamic allocation). Verasco establishes the absence of run-time errors in the analyzed programs. It enjoys a modular architecture that supports the extensible combination of multiple abstract domains, both relational and non-relational. Verasco integrates with the CompCert formally-verified C compiler so that not only the soundness of the analysis results is guaranteed with mathematical certitude, but also the fact that these guarantees carry over to the compiled code.

6.2.3. *Certified Analyses for binary codes*

Participants: Sandrine Blazy, Vincent Laporte, David Pichardie.

Static analysis of binary code is challenging for several reasons. In particular, standard static analysis techniques operate over control flow graphs, which are not available when dealing with self-modifying programs which can modify their own code at runtime. We formalized in the Coq proof assistant some key abstract interpretation techniques that automatically extract memory safety properties and control flow graphs from binary code [13], and operate over a small subset of the x86 assembly. Our analyzer is formally proved correct and has been run on several self-modifying challenges, provided by Cai et al. in their PLDI 2007 paper. This an extended version of our ITP 2014 paper.

6.3. Static analysis of functional programs using tree automata and term rewriting

Participants: Thomas Genet, Yann Salmon.

We develop a specific theory and the related tools for analyzing programs whose semantics is defined using term rewriting systems. The analysis principle is based on regular approximations of infinite sets of terms reachable by rewriting. The tools we develop use, so-called, Tree Automata Completion to compute a tree automaton recognizing a superset of all reachable terms. This over-approximation is then used to prove properties on the program by showing that some “bad” terms, encoding dangerous or problematic configurations, are not in the superset and thus not reachable. This is a specific form of, so-called, Regular Tree Model Checking. Now, we aim at applying this technique to the static analysis of programming languages whose semantics is based on terms, like functional programming languages. We already shown that static analysis of first order functional programs with a call-by-value evaluation strategy can be automated using tree automata completion [22]. This is the subject of the PhD thesis Yann Salmon has defended [11]. Now, one of the objective is to lift those results to the static analysis of higher-order functions.

6.4. Static analysis of functional specifications

Participants: Thomas Jensen, Oana Andreescu.

We have developed a static dependency analysis for a strongly typed, high-level functional specifications written in a specification formalism developed by the SME Prove & Run. In the context of interactive formal verification of complex systems, much effort is spent on proving the preservation of the system invariants. However, most operations have a localized effect on the system, which only really impacts few invariants at the same time. Identifying those invariants that are unaffected by an operation can substantially ease the proof burden for the programmer. Our dependency analysis computes a conservative approximation of the input fragments on which the operations depend. It is a flow-sensitive interprocedural analysis that handles arrays, structures and variant data types. We have validated the scalability of the analysis to complex transition systems by applying it to a functional specification of the MINIX operating system. This work was reported in [25].

6.5. Semantics

6.5.1. Energy-valued semantics

Participant: David Cachera.

We develop a $*$ -continuous Kleene ω -algebra of real-time energy functions [36]. Together with corresponding automata, these can be used to model systems which can consume and regain energy (or other types of resources) depending on available time. Using recent results on $*$ -continuous Kleene ω -algebras and computability of certain manipulations on real-time energy functions, it follows that reachability and Büchi acceptance in real-time energy automata can be decided in a static way which only involves manipulations of real-time energy functions. This works opens the way to static analysis techniques for energy-valued semantics of real-time systems.

COMETE Project-Team

7. New Results

7.1. Foundations of information hiding

Information hiding refers to the problem of protecting private information while performing certain tasks or interactions, and trying to avoid that an adversary can infer such information. This is one of the main areas of research in Comète; we are exploring several topics, described below.

7.1.1. *On the information leakage of differentially-private mechanisms*

Differential privacy aims at protecting the privacy of participants in statistical databases. Roughly, a mechanism satisfies differential privacy if the presence or value of a single individual in the database does not significantly change the likelihood of obtaining a certain answer to any statistical query posed by a data analyst. Differentially-private mechanisms are often oblivious: first the query is processed on the database to produce a true answer, and then this answer is adequately randomized before being reported to the data analyst. Ideally, a mechanism should minimize leakage, i.e., obfuscate as much as possible the link between reported answers and individuals' data, while maximizing utility, i.e., report answers as similar as possible to the true ones. These two goals, however, are in conflict with each other, thus imposing a trade-off between privacy and utility.

In [12] we used quantitative information flow principles to analyze leakage and utility in oblivious differentially-private mechanisms. We introduced a technique that exploits graph symmetries of the adjacency relation on databases to derive bounds on the min-entropy leakage of the mechanism. We considered a notion of utility based on identity gain functions, which is closely related to min-entropy leakage, and we derived bounds for it. Finally, given some graph symmetries, we provided a mechanism that maximizes utility while preserving the required level of differential privacy.

7.1.2. *Geo-indistinguishability: A Principled Approach to Location Privacy*

With the increasing popularity of handheld devices, location-based applications and services have access to accurate and real-time location information, raising serious privacy concerns for their users. In [17] we reported on our ongoing project aimed at protecting the privacy of the user when dealing with location-based services. The starting point of our approach is the principle of geo-indistinguishability, a formal notion of privacy that protects the user's exact location, while allowing approximate information – typically needed to obtain a certain desired service – to be released. We then presented two mechanisms for achieving geo-indistinguishability, one generic to sanitize locations in any setting with reasonable utility, the other custom-built for a limited set of locations but providing optimal utility. Finally we extended our mechanisms to the case of location traces, where the user releases his location repeatedly along the day and we provide a method to limit the degradation of the privacy guarantees due to the correlation between the points. All the mechanisms were tested on real datasets and compared both among themselves and with respect to the state of the art in the field.

7.1.3. *Constructing elastic distinguishability metrics for location privacy*

The recently introduced notion of geo-indistinguishability tries to address the problem of accessing location-aware services in a privacy-friendly way by adapting the well-known concept of differential privacy to the area of location-based systems. Although geo-indistinguishability presents various appealing aspects, it has the problem of treating space in a uniform way, imposing the addition of the same amount of noise everywhere on the map.

In [13] we proposed a novel elastic distinguishability metric that warps the geometrical distance, capturing the different degrees of density of each area. As a consequence, the obtained mechanism adapts the level of noise while achieving the same degree of privacy everywhere. We also showed how such an elastic metric can easily incorporate the concept of a "geographic fence" that is commonly employed to protect the highly recurrent locations of a user, such as his home or work. We performed an extensive evaluation of our technique by building an elastic metric for Paris' wide metropolitan area, using semantic information from the OpenStreetMap database. We compared the resulting mechanism against the Planar Laplace mechanism satisfying standard geo-indistinguishability, using two real-world datasets from the Gowalla and Brightkite location-based social networks. The results showed that the elastic mechanism adapts well to the semantics of each area, adjusting the noise as we move outside the city center, hence offering better overall privacy.

7.1.4. Quantitative Information Flow for Scheduler-Dependent Systems

Quantitative information flow analyses measure how much information on secrets is leaked by publicly observable outputs. One area of interest is to quantify and estimate the information leakage of composed systems. Prior work has focused on running disjoint component systems in parallel and reasoning about the leakage compositionally, but has not explored how the component systems are run in parallel or how the leakage of composed systems can be minimised.

In [23] we considered the manner in which parallel systems can be combined or scheduled. This considers the effects of scheduling channels where resources may be shared, or whether the outputs may be incrementally observed. We also generalised the attacker's capability, of observing outputs of the system, to consider attackers who may be imperfect in their observations, e.g. when outputs may be confused with one another, or when assessing the time taken for an output to appear. Our main contribution was to present how scheduling and observation affect information leakage properties. In particular, that scheduling can hide some leaked information from perfect observers, while some scheduling may reveal secret information that is hidden to imperfect observers. In addition we presented an algorithm to construct a scheduler that minimises the min-entropy leakage and min-capacity in the presence of any observer.

7.2. Foundations of Concurrency

Distributed systems have changed substantially in the recent past with the advent of phenomena like social networks and cloud computing. In the previous incarnation of distributed computing the emphasis was on consistency, fault tolerance, resource management and related topics; these were all characterized by *interaction between processes*. Research proceeded along two lines: the algorithmic side which dominated the Principles Of Distributed Computing conferences and the more process algebraic approach epitomized by CONCUR where the emphasis was on developing compositional reasoning principles. What marks the new era of distributed systems is an emphasis on managing access to information to a much greater degree than before.

7.2.1. An Algebraic View of Space/Belief and Extrusion/Utterance for Concurrency/Epistemic Logic

The notion of constraint system (cs) is central to declarative formalisms from concurrency theory such as process calculi for concurrent constraint programming (ccp). Constraint systems are often represented as lattices: their elements, called constraints, represent partial information and their order corresponds to entailment. Recently a notion of n-agent spatial cs was introduced to represent information in concurrent constraint programs for spatially distributed multi-agent systems. From a computational point of view a spatial constraint system can be used to specify partial information holding in a given agent's space (local information). From an epistemic point of view a spatial cs can be used to specify information that a given agent considers true (beliefs). Spatial constraint systems, however, do not provide a mechanism for specifying the mobility of information/processes from one space to another. Information mobility is a fundamental aspect of concurrent systems.

In the poster paper [24] we discussed using constraint systems with an algebraic operator that correspond to moving information in-between spaces as to mimic the mobility of data of distributed systems such as posting opinions/lies to other spaces or publicly disclosing data. In the conference paper [22] we enriched spatial constraint systems with operators to specify information and processes moving from a space to another. We referred to these new structures as spatial constraint systems with extrusion. We investigated the properties of this new family of constraint systems and illustrated their applications. From a computational point of view the new operators provide for process/information extrusion, a central concept in formalisms for mobile communication. From an epistemic point of view extrusion corresponds to a notion we called utterance; a piece of information that an agent communicates to others but that may be inconsistent with the agent's beliefs. Utterances can then be used to express instances of epistemic notions, which are commonplace in social media, such as hoaxes or intentional lies. Spatial constraint systems with extrusion can be seen as complete Heyting algebras equipped with maps to account for spatial and epistemic specifications. In the journal paper [28] we extended our work in [22] by showing that spatial constraint systems can also express the epistemic notion of knowledge by means of a derived spatial operator that specifies global information.

7.2.2. A Labelled Semantics for Soft Concurrent Constraint Programming

In [21] we presented a labelled semantics for Soft Concurrent Constraint Programming (SCCP), a language where concurrent agents may synchronize on a shared store by either posting or checking the satisfaction of (soft) constraints. SCCP generalizes the classical formalism by parametrising the constraint system over an order-enriched monoid: the monoid operator is not required to be idempotent, thus adding the same information several times may change the store. The novel operational rules are shown to offer a sound and complete co-inductive technique to prove the original equivalence over the unlabelled semantics.

7.2.3. Verification methods for concurrent Constraint Programming

Concurrent Constraint Programming (CCP) is a well-established declarative framework from concurrency theory. Its foundations and principles e.g., semantics, proof systems, axiomatizations, have been thoroughly studied for over the last two decades. In contrast, the development of algorithms and automatic verification procedures for CCP have hitherto been far too little considered.

To the best of our knowledge there is only one existing verification algorithm for the standard notion of CCP program (observational) equivalence. In [16] we first showed that this verification algorithm has an exponential-time complexity even for programs from a representative sub-language of CCP; the summation-free fragment (CCP+). We then significantly improved on the complexity of this algorithm by providing two alternative polynomial-time decision procedures for CCP+ program equivalence. Each of these two procedures has an advantage over the other. One has a better time complexity. The other can be easily adapted for the full language of CCP to produce significant state space reductions. The relevance of both procedures derives from the importance of CCP+. This fragment, which has been the subject of many theoretical studies, has strong ties to first-order logic and an elegant denotational semantics, and it can be used to model real-world situations. Its most distinctive feature is that of confluence, a property we exploit to obtain our polynomial procedures.

Bisimilarity is a standard behavioral equivalence in concurrency theory. However, only recently a well-behaved notion of bisimilarity for CCP, and a CCP partition refinement algorithm for deciding the strong version of this equivalence have been proposed. Weak bisimilarity is a central behavioral equivalence in process calculi and it is obtained from the strong case by taking into account only the actions that are observable in the system. Typically, the standard partition refinement can also be used for deciding weak bisimilarity simply by using Milner's reduction from weak to strong bisimilarity; a technique referred to as saturation. In [15] we demonstrated that, because of its involved labeled transitions, the above-mentioned saturation technique does not work for CCP. We also gave an alternative reduction from weak CCP bisimilarity to the strong one that allows us to use the CCP partition refinement algorithm for deciding this equivalence. We also proved that due to distinctive nature of CCP, the new method does not introduce infinitely-branching in the resulting transition systems. Finally, we derived an algorithm to automatically verify the notion of weak bisimilarity in CCP.

The ntcc calculus extends CCP with the notion of discrete time-units for the specification of reactive systems. Moreover, ntcc features constructors for non-deterministic choices and asynchronous behavior, thus allowing for (1) synchronization of processes via constraint entailment during a time-unit and (2) synchronization of processes along time-intervals. In [20] we developed the techniques needed for the automatic verification of ntcc programs based on symbolic model checking. We showed that the internal transition relation, modeling the behavior of processes during a time-unit (1 above), could be symbolically represented by formulas in a suitable fragment of linear time temporal logic. Moreover, by using standard techniques as difference decision diagrams, we provided a compact representation of these constraints. Then, relying on a fixpoint characterization of the timed constructs, we obtained a symbolic model of the observable transition (2 above). We proved that our construction is correct with respect to the operational semantics. Finally, we introduced a prototypical tool implementing our method.

COMPSYS Project-Team

7. New Results

7.1. Studying Optimal Spilling in the Light of SSA

Participants: Florian Brandner [ENSTA ParisTech, previously Compsys], Quentin Colombet [Apple, previously Compsys], Alain Darté.

Recent developments in register allocation, mostly linked to static single assignment (SSA) form, have shown the benefits of decoupling the problem in two phases: a first spilling phase places load and store instructions so that the register pressure at all program points is small enough, and a second assignment and coalescing phase maps the variables to physical registers and reduces the number of move instructions among registers. We focused on the first phase, for which many open questions remain: in particular, we studied the notion of optimal spilling (what can be expressed?) and the impact of SSA form (does it help?).

To identify the important features for optimal spilling on load-store architectures, we developed a new integer linear programming formulation, more accurate and expressive than previous approaches. Among other features, we can express SSA ϕ -functions, memory-to-memory copies, and the fact that a value can be stored simultaneously in a register and in memory. Based on this formulation, we presented a thorough analysis of the results obtained for the SPECINT 2000 and EEMBC 1.1 benchmarks, from which we have drawn, among others, the following conclusions: (1) rematerialization is extremely important; (2) SSA complicates the formulation of optimal spilling, especially because of memory coalescing when the code is not in conventional SSA (CSSA); (3) micro-architectural features are significant and thus have to be accounted for; and (4) significant savings can be obtained in terms of static spill costs, cache miss rates, and dynamic instruction counts.

Parts of this work were published at CASES 2011 [18]. The journal publication [1] contains more detailed discussions, more examples illustrating new concepts and existing approaches, and additional experiments covering the observed worst-case behavior, a new post-latency heuristic, and empiric evidence showing why static spill costs are a poor metric. Three configurations were added: Appel and George under SSA, Koes and Goldstein, and the heuristic of Braun and Hack.

7.2. Symbolic Range of Pointers in C programs

Participants: Vitor Paisante [Univ. Mineas Gerais, Brazil], Maroua Maalej, Leonardo Barbosa [Univ. Mineas Gerais, Brazil], Laure Gonnord, Fernando Pereira [Univ. Mineas Gerais, Brazil].

Alias analysis is one of the most fundamental techniques that compilers use to optimize languages with pointers. However, in spite of all the attention that this topic has received, the current state-of-the-art approaches inside compilers still face challenges regarding precision and speed. In particular, pointer arithmetic, a key feature in C and C++, is yet to be handled satisfactorily. We designed a new alias analysis algorithm to solve this problem. The key insight of our approach is to combine alias analysis with symbolic range analysis. This combination lets us disambiguate fields within arrays and structs, effectively achieving more precision than traditional algorithms. To validate our technique, we have implemented it on top of the LLVM compiler. Tests on a vast suite of benchmarks show that we can disambiguate several kinds of C idioms that current state-of-the-art analyses cannot deal with. In particular, we can disambiguate 1.35x more queries than the alias analysis currently available in LLVM. Furthermore, our analysis is very fast: we can go over one million assembly instructions in 10 seconds.

This work has been accepted at CGO'16 [30]. An extended version of the related work is available as an Inria research report [27] and will be the basis of a journal submission.

7.3. Analyzing C Programs with Arrays

Participants: Laure Gonnord, David Monniaux [CNRS/VERIMAG].

Automatically verifying safety properties of programs is hard, and it is even harder if the program acts upon arrays or other forms of maps. Many approaches exist for verifying programs operating upon Boolean and integer values (e.g., abstract interpretation, counter-examples guided abstraction refinement using interpolants), but transposing them to array properties has been fraught with difficulties.

In contrast to most preceding approaches, we do not introduce a new abstract domain or a new interpolation procedure for arrays. Instead, we generate an abstraction as a scalar problem and feed it to a preexisting solver. The intuition is that if there is a proof of safety of the program, it is likely that it can be expressed by elementary steps between properties involving only a small (tunable) number N of cells from the array.

Our transformed problem is expressed using Horn clauses over scalar variables, a common format with clear and unambiguous logical semantics, for which there exist several solvers. In contrast, solvers directly operating over Horn clauses with arrays are still very immature.

An important characteristic of our encoding is that it creates a non-linear Horn problem, with tree unfoldings, contrary to the linear problems obtained by flatly encoding the control-graph structure. Our encoding thus cannot be expressed by encoding into another control-flow graph problem, and truly leverages the Horn clause format.

Experiments with our prototype VAPHOR (see Section 6.9) show that this approach can prove automatically the functional correctness of several classical examples of the literature, including *selection sort*, *bubble sort*, *insertion sort*, as well as examples from previous articles on array analysis.

This work is presented in a research report [28] and is currently under submission.

7.4. Termination of C Programs

Participants: Laure Gonnord, David Monniaux [CNRS/VERIMAG], Gabriel Radanne [Univ Paris 7/ PPS].

The work of Compsys on the generation of multi-dimensional ranking functions [15], through a mix of polyhedral and abstract interpretation techniques, and its implementation in the tool Rank [16], was continued by Laure Gonnord in collaboration with D. Monniaux. A complete method for synthesizing lexicographic linear ranking functions (and thus proving termination), supported by inductive invariants, was designed in the case where the transition relation of the program includes disjunctions and existentials (large block encoding of control flow).

Previous work would either synthesize a ranking function at every basic block head, not just loop headers, which reduces the scope of programs that may be proved to be terminating, or expand large block transitions including tests into (exponentially many) elementary transitions, prior to computing the ranking function, resulting in a very large global constraint system. In contrast, the new algorithm incrementally refines a global linear constraint system according to extremal counterexamples: only constraints that exclude spurious solutions are included.

Experiments with our tool Termite 6.8 show marked performance and scalability improvements compared to other systems.

This work has been published at the PLDI'15 conference [7].

7.5. Data-aware Process Networks

Participants: Christophe Alias, Alexandru Plesco [XtremLogic SAS].

High-level circuit synthesis (HLS, high-level synthesis) consists in compiling a program described in a high-level programming language (as C) to a circuit. The circuit must be as efficient as possible while using properly the resources (power consumption, silicon area, FPGA elementary units, memory accesses, etc). Although a lot of progress was achieved on the back-end (low-level) aspects (pipeline generation, place/route), the front-end aspects (parallelism, I/O) are still rudimentary compared to the techniques developed by the HPC community, notably the analysis stemming from the *polyhedral model*.

We introduced data-aware process networks (DPN), a parallel execution model adapted to the hardware constraints of high-level synthesis, where the data transfers are made explicit. We have shown that the DPN model is consistent in the sense that any translation of a sequential program produces an equivalent DPN without deadlocks. Finally, we show how to compile a sequential program to a DPN and how to optimize the input/output and the parallelism.

This work has been published as an Inria research report [9] and will be submitted to a journal.

7.6. Mono-parametric Tiling

Participants: Guillaume Iooss, Sanjay Rajopadhye [Colorado State University], Christophe Alias, Yun Zou [Colorado State University].

Tiling is a crucial program transformation with many benefits. It improves locality, exposes parallelism, allows for adjusting the ops-to-bytes balance of codes, and can be applied at multiple levels. Allowing tile sizes to be symbolic parameters at compile time has many benefits, including efficient auto-tuning, and run-time adaptability to system variations. For polyhedral programs, parametric tiling in its full generality is known to be non-linear, breaking the mathematical closure properties of the polyhedral model. Most compilation tools therefore either avoid it by only performing fixed size tiling, or apply it only in the final, code generation step. Both strategies have limitations.

We first introduced mono-parametric partitioning, a restricted parametric, tiling-like transformation that can be used to express a tiling. We showed that, despite being parametric, it is a polyhedral transformation. We first proved that applying mono-parametric partitioning (i) to a polyhedron yields a union of polyhedra, and (ii) to an affine function produces a piecewise-affine function. We then used these properties to show how to partition an entire polyhedral program, including one with reductions. Next, we generalized this transformation to tiles with arbitrary tile shapes that can tessellate the iteration space (e.g., hexagonal, trapezoidal, etc). We showed how mono-parametric tiling can be applied at multiple levels, and how it enables a wide range of polyhedral analyses and transformations to be applied.

This work has been published as an Inria research report [14] and will be submitted to a journal. It is the extended version of the work published at IMPACT'14 [26].

7.7. Exact and Approximated Data-Reuse Optimizations for Tiling with Parametric Sizes

Participants: Alain Darte, Alexandre Isoard.

As mentioned in Section 7.6, loop tiling is a loop transformation widely used to improve spatial and temporal data locality, to increase computation granularity, and to enable blocking algorithms, which are particularly useful when offloading kernels on computing units with smaller memories. When caches are not available or used, data transfers and local storage must be software-managed, and some useless remote communications can be avoided by exploiting data reuse between tiles. An important parameter of tiling is the sizes of the tiles, which impact the size of the required local memory. However, for most analyses involving several tiles, which is the case for inter-tile data reuse, the tile sizes induce non-linear constraints, unless they are numerical constants. This complicates or prevents a parametric analysis with polyhedral optimization techniques.

We showed that, when tiles are executed in sequence along tile axes, the parametric (with respect to tile sizes) analysis for inter-tile data reuse is nevertheless possible, i.e., one can determine, at compile-time and in a parametric fashion, the copy-in and copy-out data sets for all tiles, with inter-tile reuse, as well as sizes for the induced local memories (this is also connected to the liveness analysis described in Section 7.12). When approximations of transfers are performed, the situation is much more complex, and involves a careful analysis to guarantee correctness when data are both read and written. We provide the mathematical foundations to make such approximations possible, thanks to the introduction of the concept of *pointwise functions*. Combined with hierarchical tiling, this result opens perspectives for the automatic generation of blocking algorithms, guided by parametric cost models, where blocks can be pipelined and/or can contain parallelism. Previous work on FPGAs and GPUs already showed the interest and feasibility of such automation with tiling, but in a non-parametric fashion.

Our method is currently implemented with the `iscc` calculator of ISL, a library for the manipulation of integer sets defined with Presburger arithmetic, a complete implementation within the PPCG compiler is in progress (see also Section 6.7).

We believe that our approximation technique can be used for other applications linked to the extension of the polyhedral model as it turns out to be fairly powerful. Our future work will be to derive efficient approximation techniques, either because the program cannot be fully analyzable, or because approximations can speed-up or simplify the results of the analysis without losing much in terms of memory transfers and/or memory sizes.

A preliminary version of this work has been presented at the IMPACT'14 workshop [19]. A revised version was published at the International Conference on Compiler Construction (CC'15) [3].

7.8. Analysis of X10 Programs

Participants: Paul Feautrier, Alain Ketterlin [Inria/CAMUS], Sanjay Rajopadhye [Colorado State University], Vijay Saraswat [IBM Research], Eric Violard [Inria/CAMUS], Tomofumi Yuki.

While, historically, Compsys has applied polyhedral analysis to sequential programs, it was recently realized that it also applies to parallel programs or specifications, with the aim of checking their correctness or improving their performance. The prospect of having to program exascale architectures, with their millions of cores, has led to the development of new programming languages, whose objective is to increase the programmer productivity. Compsys has first applied polyhedral techniques to synchronous languages [24], [25] and pipelined specifications (see Section 7.7), before concentrating on IBM's high-productivity language X10 (see this section as well as Section 7.9) and on the OpenStream language (see Section 7.10).

X10 is based on the creation of independent *activities* (light-weight threads), which can synchronize either by a generalization of the fork/join scheme, or with *clocks*, an improved version of the familiar barriers. X10 is deadlock-free by construction but it is the programmer responsibility to insure determinism by a proper use of synchronizations. Non-determinism bugs may have a very low occurrence probability thus be very difficult to detect by testing, hence the interest for detecting races at compile time. In collaboration with CSU (S. Rajopadhye, T. Yuki) and IBM (V. Saraswat), we first extended array dataflow analysis to polyhedral clock-free X10 programs [34]. We have been working on clocked programs too. Race detection becomes undecidable [35], but realistic problems may still be solved by heuristics.

In cooperation with Eric Violard and Alain Ketterlin (Inria Team Camus, Strasbourg), and in order to obtain a more secure and precise analysis, we are currently attempting to formalize the "happens before" analysis used in these two previous papers [34], [35], using the proof assistant Coq.

7.9. Revisiting Loop Transformations with X10 Clocks

Participant: Tomofumi Yuki.

Loop transformations are known to be important for performance of compute-intensive programs, and are often used to expose parallelism. However, many transformations involving loops often obfuscate the code, and are cumbersome to apply by hand. In this work, we explored alternative methods for expressing parallelism that are more friendly to the programmer. In particular, we seek to expose parallelism without significantly changing the original loop structure. We illustrated how clocks in X10 can be used to express some of the traditional loop transformations, in the presence of parallelism, in a manner that we believe to be less invasive. Specifically, expressing parallelism corresponding to one-dimensional affine schedules can be achieved without modifying the original loop structure and/or statements.

This work was published at the international workshop on X10 [8].

7.10. Static Analysis of OpenStream Programs

Participants: Albert Cohen [Inria Parkas team], Alain Darte, Paul Feautrier.

In the context of the ManycoreLabs project (see Section 8.1), we also studied the applicability of polyhedral techniques to the parallel language OpenStream [31]. When applicable, polyhedral techniques are indeed invaluable for compile-time debugging and for generating efficient code well suited to a target architecture. OpenStream is a two-level language in which a control program directs the initialization of parallel task instances that communicate through *streams*, with possibly multiple writers and readers. It has a fairly complex semantics in its most general setting, but we restricted ourselves to the case where the control program is sequential, which is representative of the majority of the OpenStream applications.

In contrast to X10, this restriction offers deterministic concurrency by construction, but deadlocks are still possible. We showed that, if the control program is polyhedral, one may statically compute, for each task instance, the read and write indices to each of its streams, and thus reason statically about the dependences among task instances (the only scheduling constraints in this polyhedral subset). If the control program has nested loops, communications use one-dimensional channels in a form of linearization, and these indices may be polynomials of arbitrary degree, thus requiring to extend to polynomials the standard polyhedral techniques for dependence analysis, scheduling, and deadlock detection. Modern SMT allow to solve polynomial problems, albeit with no guarantee of success; the approach previously developed by P. Feautrier [6] may offer an alternative solution.

The usual way of disproving deadlocks is by exhibiting a schedule for the program operations, a well-known problem for polyhedral programs where dependences can be described by affine constraints. In the case of OpenStream, we established two important results related to deadlocks: 1) a characterization of deadlocks in terms of dependence paths, which implies that streams can be safely bounded as soon as a schedule exists with such sizes, 2) the proof that deadlock detection is undecidable, even for polyhedral OpenStream.

Details of this work are available in a research report [10]. It will be presented at the international workshop IMPACT'16 [2]. Some further developments are in progress for scheduling OpenStream programs using polynomial techniques, see Section 6.4.

7.11. Handling Polynomials for Program Analysis and Transformation

Participant: Paul Feautrier.

As shown in Section 7.10, many problems in parallel programs analysis and verification can be reduced to proving or disproving properties of polynomials in the variables of the program. For instance, the so-called “linearizations” (replacing a multi-dimensional object by a one-dimensional one) generate polynomial access functions. These polynomials then reappear in dependence testing, scheduling, and invariant construction. It may also happen that polynomials are absent from the source program, but are created either by an enabling analysis, as for OpenStream, or are imposed by complexity consideration. The usual solution is to construct a multi-dimensional function (e.g., a schedule for parallelization or a ranking function for termination [15]), which can then be converted into polynomials by counting. However, a direct approach is preferable, especially when the resulting schedule is to be used for further analysis, e.g., in real-time situations or WCET evaluation.

What is needed here is a replacement for the familiar emptiness tests and for Farkas lemma (deciding whether an affine form is positive inside a polyhedron). Recent mathematical results by Handelman and Schweighofer on the *Positivstellensatz* allow one to devise algorithms that are able to solve these problems. The difference is that one gets only sufficient conditions, and that complexity is much higher than in the affine cases. A paper presenting applications of these ideas to three use cases – dependence testing, scheduling, and transitive closure approximation – was presented at the 5th International Workshop on Polyhedral Compilation Techniques (IMPACT'15) [6] in Amsterdam in January 2015. A tool implementing polyhedral schedules complements this work, see Section 6.6.

7.12. Liveness Analysis in Explicitly-Parallel Programs

Participants: Alain Darte, Alexandre Isoard, Tomofumi Yuki.

In the light of the parallel specifications encountered in our other works (from Section 7.7 to Section 7.11), we revisited scalar and array element-wise liveness analysis for programs with parallel specifications. In earlier work on memory allocation/contraction (register allocation or intra- and inter-array reuse in the polyhedral model), a notion of “time” or a total order among the iteration points was used to compute the liveness of values. In general, the execution of parallel programs is not a total order, and hence the notion of time is not applicable.

We first revised how conflicts are computed by using ideas from liveness analysis for register allocation, studying the structure of the corresponding conflict/interference graphs. Instead of considering the conflict between two pairs of live ranges, we only consider the conflict between a live range and a write. This simplifies the formulation from having four instances involved in the test down to three, and also improves the precision of the analysis in the general case.

Then we extended the liveness analysis to work with partial orders so that it can be applied to many different parallel languages/specifications with different forms of parallelism. An important result is that the complement of the conflict graph with partial orders is directly connected to memory reuse, even in presence of races. However, programs with conditionals do not even have a partial order, and our next step will be to handle such cases with more accuracy.

Details of this work are available in a research report [13]. It will be presented at the international workshop IMPACT’16 [4].

7.13. Extended Lattice-Based Memory Allocation

Participants: Alain Darte, Alexandre Isoard, Tomofumi Yuki.

We extended lattice-based memory allocation [20], an earlier work on memory (array) reuse analysis. The main motivation is to handle in a better way the more general forms of specifications we see today, e.g., with loop tiling, pipelining, and other forms of parallelism available in explicitly parallel languages. Our extension has two complementary aspects. We showed how to handle more general specifications where conflicting constraints (those that describe the array indices that cannot share the same location) are specified as a (non-convex) union of polyhedra. Unlike convex specifications, this also requires to be able to choose suitable directions (or basis) of array reuse. For that, we extended two dual approaches, previously proposed for a fixed basis, into optimization schemes to select suitable basis. Our final approach relies on a combination of the two, also revealing their links with, on one hand, the construction of multi-dimensional schedules for parallelism and tiling (but with a fundamental difference that we identify) and, on the other hand, the construction of universal reuse vectors (UOV), which was only used so far in a specific context, for schedule-independent mapping.

This algorithmic work, connected to the parametric tiling of Section 7.7 and the liveness analysis results of Section 7.12, is complemented by a set of prototype scripting tools, see Section 6.3.

Details of this work are available in a research report. It has also been submitted to a conference.

7.14. Stencil Accelerators

Participants: Steven Derrien [University of Rennes 1, Inria/CAIRN], Xinyu Niu [Imperial College London], Sanjay Rajopadhye [Colorado State University], Tomofumi Yuki.

Stencil computations have been known to be an important class of programs for scientific calculations. Recently, various architectures (mostly targeting FPGAs) for stencils are being proposed as hardware accelerators with high throughput and/or high energy efficiency. There are many different challenges for such design: How to maximize compute-I/O ratio? How to partition the problem so that the data fits on the on-chip memory? How to efficiently pipeline? How to control the area usage? We seek to address these challenges by combining techniques from compilers and high-level synthesis tools.

One project in collaboration with the CAIRN team and Colorado State University targets stencils with regular dependence patterns. Although many architectures have been proposed for this type of stencils, most of them use a large number of small processing elements (PE) to achieve high throughput. We are exploring an alternative design that aims for a single, large, deeply-pipelined PE. The hypothesis is that the pipelined parallelism is more area-efficient compared to replicating small PEs. We have published a work-in-progress paper on this topic at IMPACT'16 [5].

Another type of stencil accelerators that we are working on, in collaboration with Xinyu Niu, targets stencil programs with dynamic dependences (i.e., sparse computations). The collaboration is in the context of the EURECA project ⁰ where the dynamic reconfigurability of modern FPGAs are used to efficiently handle dynamic access patterns.

7.15. PolyApps

Participant: Tomofumi Yuki.

Loop transformation frameworks using the polyhedral model have gained increased attention since the rise of the multi-core era. We now have several research tools that have demonstrated their power on important kernels found in scientific computations. However, there remains a large gap between the typical kernels used to evaluate these tools and the actual applications used by the scientists.

PolyApps is an effort to collect applications from other domains of science to better establish the link between the compiler tools and “real” applications. The applications are modified to bypass some of the front-end issues of research tools, while keeping the ability to produce the original output. The goal is to assess how the state-of-the-art automatic parallelizers perform on full applications, and to identify new opportunities that only arise in larger pieces of code.

We showed that, with a few enhancements, the current tools will be able to reach and/or exceed the performance of existing parallelizations of the applications. One of the most critical element missing in current tools is the ability to modify the memory mappings.

⁰<http://www.doc.ic.ac.uk/~nx210/2015/09/01/eureca.html>

CONVECS Project-Team

6. New Results

6.1. New Formal Languages and their Implementations

6.1.1. Definition of LNT

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

LNT is a next generation formal description language for asynchronous concurrent systems, which attempts to combine the best features of imperative programming languages and value-passing process algebras. LNT is increasingly used by CONVECS for industrial case studies and applications (see § 6.5) and serves also in university courses on concurrency, in particular at ENSIMAG (Grenoble) and at Saarland University.

In 2015, the theoretical foundations of LNT have been explored in a journal article [14] that, after examining the various ways sequential composition is handled in mainstream value-passing process calculi, shows that these various approaches are subsumed by the LNT approach, which is easier to learn and leads to more readable and more concise specifications.

The LNT language has also been enhanced in several aspects:

- The “case” construct now supports multiple (tuple-like) expressions and patterns.
- Two new parameter-passing modes “in var” and “out var” have been introduced to allow finer data-flow analyses.
- Exceptions are better handled and a new “assert” statement was added to LNT.
- The “none” channel is now implicitly predefined.
- Finally, the LNT reference manual has been extended and updated at many places.

6.1.2. Translation from LNT to LOTOS

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

In 2015, the translator from LNT to LOTOS was further improved. In addition to 22 bug fixes and improved error messages, the following enhancements have been brought:

- The “-root” option of LNT2LOTOS now accepts value parameters for LNT processes and supports gate parameters in named style. It also accepts the name of a process not present in the current module.
- Negative number constants of the form “ -2^{k-1} ”, where integer numbers are represented using k bits, are now supported.
- Better warning messages are emitted for “in” and “in out” (formerly “inout”) parameters.

6.1.3. Translation from LOTOS to Petri nets and C

Participants: Hubert Garavel, Wendelin Serwe.

The LOTOS compilers CAESAR and CAESAR.ADT, which were once the flagship of CADP, now play a more discrete role since LNT (rather than LOTOS) has become the recommended specification language of CADP. Thus, CAESAR and CAESAR.ADT are mostly used as back-end translators for LOTOS programs automatically generated from LNT or other formalisms such as Fiacre, and are only modified when this appears to be strictly necessary.

In 2015, in addition to a few bug fixes, the “-root” option of the CAESAR compiler has been generalized to support processes having value parameters; this makes compositional verification easier by removing the need for introducing extra wrapper processes having no value parameters. The EXEC/CAESAR interface has been extended with two new primitives “CAESAR_KERNEL_DELAY” and “CAESAR_KERNEL_EXIT()”. Also, optimizations have been introduced to generate shorter and simpler C code, and to make sure that this C code compiles without spurious warnings.

A systematic comparison between CAESAR.ADT and available interpreters/compiler for other languages that support rewrite rules or pattern matching has been undertaken. This comparison reuses the benchmarks developed for the three Rewrite Engine Competitions (2006, 2009, and 2010). As a preliminary step, we developed a tenth translators from the REC formalism in which these benchmarks are written to languages such as Haskell, LOTOS, Maude, mCRL, OCAML, Opal, Rascal, Scala, and Tom.

6.1.4. NUPN

Participants: Hubert Garavel, Frédéric Lang.

The CAESAR.BDD tool that analyzes NUPN (*Nested-Unit Petri Nets*) models and serves to prepare the yearly Model Checking Contest ⁰ has been enhanced in several ways. In addition to 7 bug fixes, 14 new command-line options have been added to CAESAR.BDD (“-arcs”, “-bits”, “-creator”, “-density”, “-encodings”, “-height”, “-hwb”, “-multiple-arcs”, “-multiple-initial-tokens”, “-places”, “-redundant-units”, “-transitions”, “-units”, and “-width”). The output format produced by the “-exclusive-places” option has been revised. The “-mcc” option now computes the extended free choice property. A new option “-network nupn” was also added to EXP.OPEN to produce NUPN models from automata networks.

Particular efforts have been put to increase the scalability of CAESAR.BDD for large models. Reading large NUPN files was made much faster. The “-exclusive-places” option of CAESAR.BDD was made faster too. The size of the largest data structure allocated by CAESAR.BDD, has been divided by four. CAESAR.BDD has also been optimized to save memory when handling NUPN models having a simple hierarchical structure. Finally, user-specified timeouts are better supported.

A conference article was published [24], which formally defines the NUPN model and investigates its mathematical properties. Additionally, the assembly of a collection of large NUPN models was undertaken, and various prototype tools to handle NUPN models (“nupn_pack”, “nupn_reduce”, and “nupn_merge”) have been developed.

6.1.5. Translation from GRL to LNT

Participants: Fatma Jebali, Jingyan Jourdan-Lu, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the formal modeling of GALS (*Globally Asynchronous, Locally Synchronous*) systems, which are composed of several synchronous subsystems evolving cyclically, each at its own pace, and communicating with each other asynchronously. Designing GALS systems is challenging due to both the high level of (synchronous and asynchronous) concurrency and the heterogeneity of computations (deterministic and nondeterministic). To bring our formal verification techniques and tools closer to the GALS paradigm, we designed a new formal language named GRL (*GALS Representation Language*), as an intermediate format between GALS models and purely asynchronous concurrent models. GRL combines the main features of synchronous dataflow programming and asynchronous process calculi into one unified language, while keeping the syntax homogeneous for better acceptance by industrial GALS designers. GRL allows a modular composition of synchronous systems (blocks), environmental constraints (environments), and asynchronous communication mechanisms (mediums), to be described at a level of abstraction that is appropriate to verification. GRL also supports external C and LNT code. A translator named GRL2LNT has been developed, allowing an LNT program to be generated from a GRL specification automatically. Additionally, an OPEN/CAESAR-compliant compiler named GRL.OPEN (based on GRL2LNT and LNT.OPEN) makes possible the on-the-fly exploration of the LTS underlying a GRL specification using CADP.

⁰<http://mcc.lip6.fr/>

In 2015, we have revised the GRL syntax to make GRL easier to learn and to understand. Our data base of examples has been updated to take those changes into account. We have also added some language features, such as named constants, and a dedicated construct called *activation signal* to define constraints on the asynchronous activation of blocks. This new construct is easier to use than the previous solution based on ad-hoc data signals, and semantically more elegant as it avoids unexpected deadlocks. Activation signals permit realistic situations such as halting, priorities, scenarios, and pace relations between synchronous components to be modeled at a suitable level of abstraction. They can be smoothly translated into LNT without affecting the rest of the translation.

As regards the specification of properties, to reduce the complexity of using full-fledged temporal logics, we have also proposed a property specification language dedicated to GALS systems, based upon a set of temporal logic patterns, which capture frequently encountered behaviours, encompassing both time-critical and untimed aspects of GALS systems. Those patterns include deadlock, livelock, safety, liveness, and fairness properties. The semantics of the proposed patterns have been defined by translation into the MCL language.

As regards the GRL2LNT tool, nine successive versions have been released, to take into account the syntactic changes in the GRL language, to correct about 20 bugs, to eliminate compilation warnings, and to implement the following new features:

- The generated LNT code has been corrected so as to eliminate compilation warnings and to take into account recent changes in the syntax of LNT (see § 6.1.1).
- GRL system specifications can now be parameterized with data values and instantiated using the new “-root” option of GRL2LNT.
- The interface between GRL and external C code has been revised in two ways: (1) external blocks with several outputs are now mapped to a single external function instead of one function per output previously, and (2) conversion functions between GRL and C numeric types have been defined, handling runtime verification of overflows. Those conversion functions have been packaged in a new code library, which is automatically included by GRL2LNT.
- Several verifications on the usage of signals and communication channels have been implemented, leading either to error messages, or to warnings corresponding to potential errors. About 20 new error messages and 10 new warnings have been added.

In addition, three manual pages have been written to document respectively the GRL language, the GRL2LNT translator tool, and the GRL.OPEN shell script. The GRL non-regression test base has been extended and now contains 230 correct examples and 400 incorrect examples.

An article describing the GRL language and its associated tools has been submitted to an international journal.

6.1.6. Translation from BPMN to LNT

Participant: Gwen Salaün.

Business processes support the modeling and the implementation of software as workflows of local and inter-process activities. Taking over structuring and composition, evolution has become a central concern in software development. We believe this should be taken into account as soon as the modeling of business processes, which can thereafter be made executable using process engines or model-to-code transformations. We advocate that business process evolution can be formally analyzed in order to compare different versions of processes, identify precisely the differences between them, and ensure the desired consistency.

To reach this objective, we developed, in collaboration with Pascal Poizat (LIP6, Paris), a model transformation from the BPMN standard notation to the LNT process algebra. We then proposed a set of relations for comparing business processes at the formal model level. With reference to related work, we proposed a richer set of comparison primitives supporting renaming, refinement, property- and context-awareness. Thanks to the implementation of a tool that integrates with the Eclipse IDE and behind-the-scene interaction with the CADP verification toolbox, we put the checking of evolution within the reach of business process designers. Our approach is fully automated and has been applied for evaluation to a large set of BPMN processes.

6.1.7. Other Language Developments

Participants: Hugues Evrard, Hubert Garavel, Frédéric Lang, Eric Léo, Wendelin Serwe.

The ability to compile and verify formal specifications with complex, user-defined operations and data structures is a key feature of the CADP toolbox since its very origins. A long-run effort has been recently undertaken to ensure a uniform treatment of types, values, and functions across all the various CADP tools.

In 2015, the connection to external software development tools has progressed. The support of the LOTOS and LNT languages in the Emacs/XEmacs, jEdit, and Vim editors has improved. More text editors are now supported, including Nano, Notepad++, and all the text editors compliant with GtkSourceView 3.0 (including the Gedit editor of Gnome). Pretty-printers such as a2ps and the LaTeX "listings" package are also supported. Configuration files for three CADP languages (MCL, SVL, and XTL) and three CADP formats (BES, NUPN, and RBC) have been added.

Also, with the help of its principal author Pierre Boullier (Inria, Alpage), we corrected a memory allocation bug in the SYNTAX parser generator, which is used in most of the compilers developed by the CONVECS team.

6.2. Parallel and Distributed Verification

6.2.1. Distributed Code Generation for LNT

Participants: Hugues Evrard, Frédéric Lang.

Rigorous development and prototyping of a distributed algorithm using LNT involves the automatic generation of a distributed implementation. For the latter, a protocol realizing process synchronization is required. As far as possible, this protocol must itself be distributed, so as to avoid the bottleneck that would inevitably arise if a unique process would have to manage all synchronizations in the system. A particularity of such a protocol is its ability to support branching synchronizations, corresponding to situations where a process may offer a choice of synchronizing actions (which themselves may nondeterministically involve several sets of synchronizing processes) instead of a single one. Therefore, a classical barrier protocol is not sufficient and a more elaborate synchronization protocol is needed.

Using a synchronization protocol that we verified formally in 2013, we developed a prototype distributed code generator, named DLC (*Distributed LNT Compiler*), which takes as input the model of a distributed system described as a parallel composition of LNT processes.

In 2015, we finalized the development of DLC: the code was cleaned and the different compiler components were better integrated. A new option was added for the generated executables to dump at runtime an execution trace in the SEQUENCE format of CADP, for further analysis. A complete description of DLC, its synchronization protocol, performance data and usage examples were presented in Hugues Evrard's PhD thesis [9], defended in July 2015. An overview of DLC was presented in an international conference paper [23], and an extended version has been prepared for a journal special issue currently under construction. A tool paper was accepted in an international conference that will take place in 2016 [22].

6.2.2. Verification of Asynchronously Communicating Systems

Participants: Lakhdar Akroun, Gwen Salaün.

Analyzing systems communicating asynchronously via reliable FIFO buffers is an undecidable problem. A typical approach is to check whether the system is bounded, and if not, whether the corresponding state space can be made finite by limiting the presence of communication cycles in behavioral models or by fixing the buffer size. In this work, our focus is on systems that are likely to be unbounded and therefore result in infinite systems. We do not want to restrict the system by imposing any arbitrary bound. We introduced a notion of stability and proved that once the system is stable for a specific buffer bound, it remains stable whatever larger bounds are chosen for buffers. This enables one to check certain properties on the system for that bound and to ensure that the system will preserve them whatever larger bounds are used for buffers. We also proved that computing this bound is undecidable but we showed how we can succeed in computing these bounds for many typical examples using heuristics and equivalence checking.

6.2.3. Analysis of Verification Counterexamples

Participants: Gianluca Barbon, Gwen Salaün.

Model checking is an established technique for automatically verifying that a model, e.g., a Labelled Transition System (LTS), obtained from higher-level specification languages (such as process algebras) satisfies a given temporal property, e.g., the absence of deadlocks. When the model violates the property, the model checker returns a counterexample, which is a sequence of actions leading to a state where the property is not satisfied. Understanding this counterexample for debugging the specification is a complicated task for several reasons: (i) the counterexample can contain hundreds (even thousands) of actions, (ii) the debugging task is mostly achieved manually, and (iii) the counterexample does not give any clue on the state of the system (e.g., parallelism or data expressions) when the error occurs.

In collaboration with the SLIDE team of the LIG laboratory, we work on new solutions for simplifying the comprehension of counterexamples and thus favouring usability of model checking techniques. To do so, we apply pattern mining techniques to a set of correct traces (extracted from the LTS) and incorrect traces (corresponding to counterexamples), to identify specific patterns indicating more precisely the source of the problem.

6.3. Timed, Probabilistic, and Stochastic Extensions

6.3.1. Model Checking for Extended PCTL

Participants: Radu Mateescu, José Ignacio Requeno.

In the context of the SENSATION project (see § 8.2.1.1), we study the specification and verification of quantitative properties of concurrent systems.

In 2015, we developed a probabilistic version of ACTL (Action-based CTL) [41], named PACTL. This logic represents an action-based counterpart for PCTL (*Probabilistic Computation Tree Logic*) [50] and is interpreted naturally over DTMCs with labeled transitions, such as those produced from IPCs (*Interactive Probabilistic Chains*) [40]. The PACTL operators generalize those of ACTL: they characterize sequences of transitions in the DTMC by specifying both the states and the actions labeling the transitions. We implemented PACTL as an XTL library, which allows the designer to combine properties on actions, data, probabilities, and discrete time. We have experimented the PACTL library on several DTMCs imported from the probabilistic model checker PRISM [55] to ensure that both implementations provide the same numerical results.

6.4. Component-Based Architectures for On-the-Fly Verification

6.4.1. Compositional Verification

Participants: Hubert Garavel, Frédéric Lang.

The CADP toolbox contains various tools dedicated to compositional verification, among which EXP.OPEN, BCG_MIN, BCG_CMP, and SVL play a central role. EXP.OPEN explores on the fly the graph corresponding to a network of communicating automata (represented as a set of BCG files). BCG_MIN and BCG_CMP respectively minimize and compare behavior graphs modulo strong or branching bisimulation and their stochastic extensions. SVL (*Script Verification Language*) is both a high-level language for expressing complex verification scenarios and a compiler dedicated to this language.

In 2015, we corrected one bug in BCG_CMP and eight bugs in SVL. We extended the SVL language and compiler as follows:

- A new statement was added to translate a LOTOS file or a process in a LOTOS file to an LNT file automatically.
- LNT processes with data parameters can now be instantiated directly in the SVL script, without requiring a parameterless intermediate process to be defined.

- LNT processes with gate parameters can now be instantiated in the SVL script using the named parameter-passing style of LNT.
- Specification of a diagnostic file is now optional in the “comparison”, “deadlock”, and “live-lock” statements of SVL.
- The “property” statement has been extended so that it can now contain any kind of statement, provided it contains at least one verification statement.
- Within SVL properties, it is now possible to define shell lines followed by an “expected” clause to specify the expected result of the shell line.
- It is now possible to add a “result” clause after a verification statement, so as to store the result of the verification in a shell variable that can be subsequently used in the SVL script.

We improved several demo examples of CADP by using these new SVL constructs, and we added a new demo example on the verification of an airplane-ground communication protocol.

We also improved the PMC tool, by correcting five bugs and adding a new “-order” option, which permits the user to define a particular order for quotienting. We improved the presentation of the demo examples released in the PMC distribution. Those examples are now given in LNT and translated automatically into networks of automata in the EXP language, instead of being given directly as networks of automata.

6.4.2. On-the-Fly Test Generation

Participants: Hubert Garavel, Radu Mateescu, Wendelin Serwe.

In the context of the collaboration with STMicroelectronics, we study techniques for testing if a (hardware) implementation is conform to a formal model described in LNT. Our approach is inspired by the theory of conformance testing [63], as implemented for instance in TGV [53] and JTorX [33]. We have developed three prototype tools to support this approach. The first tool implements a dedicated OPEN/CAESAR-compliant compiler for the particular asymmetric synchronous product between the model and the test purpose. The second tool, based on slightly extended generic components for graph manipulation (τ -compression, τ -confluence reduction, determinization) and resolution of Boolean equation systems, generates the complete test graph (CTG), which can be used to extract concrete test cases or to drive the test of the implementation. A third prototype tool takes as input a CTG and extracts either a single test case (randomly chosen or the first encountered one), or the set of *all* test cases. The principal advantage of our approach compared to existing tools is the use of LNT for describing test purposes, which facilitates the manipulation of data values.

In 2015, we corrected the prototype tools to properly handle timers and failure transitions, improved the documentation, and simplified internal data structures.

These prototype tools were used in the case study with STMicroelectronics (see § 6.5.1) and the EnergyBus (see § 6.5.4).

6.4.3. Other Component Developments

Participants: Soraya Arias, Hubert Garavel, Frédéric Lang, Radu Mateescu.

We separated the MCL library defining the operators of ACTL (Action-based CTL) [41] in two parts: the first one defines the operators of $ACTL \setminus X$ (the fragment of ACTL without the next-time operators), including optimized definitions of derived temporal operators, and the second one defines the next-time operators, including the definitions of silent next-time operators, which complement the visible next-time operators already present in the library.

We also added an MCL library defining the operators of the $L\mu$ -dsbr fragment of modal μ -calculus [6], which includes the $ACTL \setminus X$ library. The $L\mu$ -dsbr library also defines the absence of deadlock property as an MCL formula adequate w.r.t. divergence-sensitive branching bisimulation (divbranching for short) and allowing one to hide all visible actions in the LTS and to reduce it modulo divbranching prior to verification, which may bring significant performance gains.

A new major version 1.2 of the BCG format for storing Labelled Transition Systems was released as part of CADP 2015-a. Following this change, various minor residual bugs have been identified and fixed in 2015, and the type system of XTL has been modified to require fewer explicit type coercions.

In addition to bug fixes in various tools (e.g., CUNCTATOR, EUCALYPTUS, TST, XTL, etc.), the installation procedures of CADP have been revisited and updated; in particular, work is going on and many preliminary changes have been silently brought to ease installation of CADP on Windows.

6.5. Real-Life Applications and Case Studies

6.5.1. ACE Cache Coherency Protocol

Participants: Abderahman Kriouile, Radu Mateescu, Wendelin Serwe.

In the context of a CIFRE convention with STMicroelectronics, we studied system-level cache coherency, a major challenge faced in the current System-on-Chip architectures. Because of their increasing complexity (mainly due to the significant number of computing units), the validation effort using current simulation-based techniques grows exponentially. As an alternative, we study formal verification.

We focused on the ACE (AXI Coherency Extensions) cache coherency protocol, a system-level coherency protocol proposed by ARM [31]. In previous years, we developed a parametric formal model (about 3,400 lines of LNT) of a system consisting of an ACE-based cache coherent interconnect, processors, and a main memory. We also specified temporal properties expressing cache coherence, data integrity, and successful completion of each transaction. Note that the former property required to transform state-based properties into action-based ones, by adding information about the cache state to the actions executed by the cache.

In 2015, we continued to exploit the formal model to improve the validation of the architecture under design at STMicroelectronics, in particular by integrating tests derived from the formal model into the official test plans. This work led to a publication in an international conference [25], and the defense of the PhD corresponding to the CIFRE convention [10].

6.5.2. Deployment and Reconfiguration Protocols for Cloud Applications

Participants: Rim Sakka Abid, Gwen Salaün.

Cloud applications are complex applications composed of a set of interconnected software components running on different virtual machines, hosted on remote physical servers. Deploying and reconfiguring this kind of applications are very complicated tasks especially when one or multiple virtual machines fail when achieving these tasks. Hence, there is a need for protocols that can dynamically reconfigure and manage running distributed applications.

In 2015, we proposed a novel protocol, which aims at reconfiguring cloud applications. This protocol is able to ensure communication between virtual machines and resolve dependencies by exchanging messages, (dis)connecting, and starting/stopping components in a specific order. The interaction between machines is assured via a publish-subscribe messaging system. Each machine reconfigures itself in a decentralized way. The protocol supports virtual machine failures, and the reconfiguration always terminates successfully even in the presence of a finite number of failures. Due to the high degree of parallelism inherent to these applications, the protocol was specified in LNT and verified using CADP. The use of formal specification languages and tools helped to detect several bugs and to improve the protocol. These results were published in [12].

Another line of work concerns autonomic computing in cloud environments. Managing distributed cloud applications is a challenging problem because manual administration is no longer realistic for these complex distributed systems. Thus, autonomic computing is a promising solution for monitoring and updating these applications automatically. This is achieved through the automation of administration functions and the use of control loops called autonomic managers. An autonomic manager observes the environment, detects changes, and reconfigures dynamically the application. Multiple autonomic managers can be deployed in the same system and must make consistent decisions. Using them without coordination may lead to inconsistencies and error-prone situations.

In 2015, we propose an approach for coordinating stateful autonomic managers, which relies on a simple coordination language, new techniques for asynchronous controller synthesis and Java code generation. We used our approach for coordinating real-world cloud applications. These results were published in [19].

6.5.3. Networks of Programmable Logic Controllers

Participants: Fatma Jebali, Jingyan Jourdan-Lu, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the software applications embedded on the PLCs (Programmable Logic Controllers) manufactured by Crouzet Automatismes. One of the objectives of Bluesky is to enable the rigorous design of complex control applications running on several PLCs connected by a network. Such applications are instances of GALS (*Globally Asynchronous, Locally Synchronous*) systems composed of several synchronous automata embedded on individual PLCs, which interact asynchronously by exchanging messages. A formal analysis of these systems can be naturally achieved by using the formal languages and verification techniques developed in the field of asynchronous concurrency.

For describing the applications embedded on individual PLCs, Crouzet provides a dataflow language with graphical syntax and synchronous semantics, equipped with an ergonomic user-interface that facilitates the learning and use of the language by non-experts. To equip the PLC language of Crouzet with functionalities for automated verification, the solution adopted in Bluesky was to translate it into GRL (see § 6.1.5), which enables the connection to testing and verification tools covering the synchronous and asynchronous aspects.

In 2015, we have provided support to Crouzet, who started to integrate GRL in the PLC design process by developing both a library of GRL blocks corresponding to function blocks present in their PLC programming tool, and an automated translation from a PLC program into a GRL block. The GRL2LNT and GRL.OPEN tools (see § 6.1.5) provide a direct connection to all verification functionalities of CADP, in particular model checking and equivalence checking.

We also investigated the equivalence checking for networks of PLCs, with the objective of proposing a general methodology usable in industrial context. We identified several rules (formalized as templates) for describing the asynchronous and synchronous parts of PLC networks, as well as their external behaviour (service), in order to facilitate the equivalence checking modulo branching bisimulation.

6.5.4. EnergyBus Standard for Connecting Electric Components

Participants: Hubert Garavel, Wendelin Serwe.

The EnergyBus⁰ is an upcoming industrial standard for electric power transmission and management, based on the CANopen field bus. It is developed by a consortium assembling all major industrial players (such as Bosch, Panasonic, and Emtas) in the area of light electric vehicles (LEV); their intention is to ensure interoperability between all electric LEV components. At the core of this initiative is a universal plug integrating a CAN-Bus⁰ with switchable power lines. The central and innovative role of the EnergyBus is to manage the safe electricity access and distribution inside an EnergyBus network.

In the framework of the European FP7 project SENSATION (see § 8.2.1.1) a formal specification in LNT of the main EnergyBus protocols is being developed by Alexander Graf-Brill and Holger Hermanns at Saarland University [48], with the active collaboration of CONVECS.

In 2015, we pursued the analysis of the LNT model, involving both verification (by means of state-space exploration and model checking techniques) and validation (using test cases automatically derived from the formal LNT model).

6.5.5. AutoFlight Control System

Participant: Fatma Jebali.

⁰<http://www.energybus.org>

⁰<http://www.can-cia.org>

In collaboration with Eric Jenn (IRT Saint Exupery, Toulouse), we studied an AutoFlight Control System (AFCS), provided by Thales Avionics. The goal of an AFCS is to improve the quality of flight and enhance the operational capability of the aircraft. The architecture of the AFCS comprises two parts. The first part (FCP, Flight Control Panel) consists of a control panel, which enables the pilot to interact with the system. The second part (AFS, Automatic Flight System) is in charge of performing functions such as guidance and automatic pilot. For safety purposes, each part is organized into a command and monitoring channels. The command channel ensures the function allocated to the component. The monitoring channel ensures that the command channel operates correctly. To ensure a sufficient availability level, a high level of redundancy is built inside the system. Components communicate using various communication means with different latencies (AFDX, A429, discrete).

Since AFCSs have stringent safety and time-critical requirements, formal verification is required to ensure their correctness. To this aim, we have applied the GRL approach for the formal modelling and verification of GALS systems (see § 6.1.5). In a first step, we have addressed the AFCS without redundancy. We have written a GRL description (750 lines), which can be parameterized by the activation paces of different synchronous components. We have written a set of correctness properties in MCL, which we have verified on the GRL model.

6.5.6. Graphical User-Interfaces and Plasticity

Participants: Hubert Garavel, Frédéric Lang, Raquel Oliveira.

In the context of the Connexion project (see § 8.1.1.2) and in close collaboration with Gaëlle Calvary and Sophie Dupuy-Chessa (IIHM team of the LIG laboratory), we study the formal description and validation of graphical user-interfaces using the most recent features of the CADP toolbox. The case study assigned to LIG in this project is a prototype graphical user-interface [38] designed to provide human operators with an overview of a running nuclear plant. The main goal of the system is to inform the operators about alarms resulting from faults, disturbances, or unexpected events in the plant. Contrary to conventional control rooms, which employ large desks and dedicated hardware panels for supervision, this new-generation interface uses standard computer hardware (i.e., smaller screen(s), keyboard, and mouse), thus raising challenging questions on how to best provide synthetic views of the plant status. Another challenge is to introduce plasticity in such interface, so as to enable several supervision operators, including mobile ones outside of the control room, to get accurate information in real time.

We formally specified this new-generation interface in LNT, encompassing not only the usual components traditionally found in graphical user-interfaces, but also a model of the physical world (namely, a nuclear reactor with various fault scenarios) and a cognitive model of a human operator in charge of supervising the plant. Also, several desirable properties of the interface have been expressed in MCL and verified on the LNT model using CADP. At last, we used our formal model to check conformance of execution traces generated by an industrial control room prototype provided by a partner in the project.

In 2015, we finalized our approach to formally verifying safety critical interactive systems provided with plastic user interfaces, either using equivalence checking (to check whether different versions of user interfaces present the same interaction capabilities and appearance) or model checking (to check a set of properties over a model of the system). The results have been published in international conferences [26], [27] and journals [17], and in Raquel Oliveira's PhD thesis [11].

6.5.7. Fault-Tolerant Routing for Network-on-Chip Architectures

Participant: Wendelin Serwe.

Fault-tolerant architectures provide adaptivity for on-chip communications, but also increase the complexity of the design, so that formal verification techniques are needed to check their correctness. In collaboration with Chris Myers and Zhen Zhang (University of Utah, USA), we studied an extension of the link-fault tolerant Network-on-Chip (NoC) architecture introduced by Wu *et al* [64] that supports multiflit wormhole routing. A major difference with similar architectures existing in the literature is that the considered routing algorithm is not statically proven free of deadlocks, but rather implements deadlock avoidance (by dynamically detecting possible deadlock situations and avoiding them by dropping packets).

In 2015, we detected a potential livelock in the previously developed formal LNT model [65]. The correction of this problem led to an improved routing algorithm, for which the state space for 2x2 NoCs could be generated compositionally. We also experimented with the analysis of larger configurations on Grid'5000, but even a 2x3 NoC is still too large, so that compositional state space generation fails with an intermediate state space of several billions of states. This work led to a publication accepted in an international journal [18] and a PhD thesis [66].

6.5.7.1. Other Case Studies

The demo examples of CADP, which have been progressively accumulated since the origins of the toolbox, are a showcase for the multiple capabilities of CADP, as well as a test bed to assess the new features of the toolbox. In 2015, the effort to maintain and enhance these demos has been pursued. The progressive migration to LNT has continued, by translating five demos (16, 21, 22, 36, and 38) from LOTOS to LNT. A new demo 05 (airplane-ground communication protocol) has been added. The code of many demos was updated to use the latest features of LNT, such as “`in var`” parameters and “`assert`” statements. Demos 14 and 16 have been greatly simplified by inlining MCL and XTL temporal logic formulas in SVL scripts, using the “`property`”, “`check`”, and “`|=`” statements recently added to SVL. Nine demos (02, 08, 17, 20, 27, 28, 31, 33, and 36) have been simplified by using the new possibility to pass value parameters to LOTOS and LNT processes directly in SVL scripts. XTL formulas have been shortened in demos 23 and 27. The illustration of the EXEC/CAESAR framework in demo 38 has been integrated as a property into the SVL script. Finally, demo 38 led to a publication in an international workshop [29].

CORSE Team

6. New Results

6.1. An interval constrained memory allocator for the Givy GAS runtime

Participants: François Gindraud, Fabrice Rastello, Albert Cohen [ENS Ulm], Francois Broquedis.

This work presents a memory allocator for a global address space (GAS) execution environment targeting manycore architectures with distributed memory. Among the family of Multi Processor System on Chip (MPSoC), these devices are composed of multiple nodes linked by an on-chip network; most nodes have multiple processors sharing a small local memory. An MPSoC has an excellent performance-per-Watt ratio, but it is hard to program due to multilevel parallelism, explicit resource and memory management, and hardware constraints (limited memory, network topology).

Practical programming frameworks let the programmer in charge of the hard, target-specific work (e.g., threads or node-local OpenMP plus explicit communications). Automatic, more abstract frameworks exist for specific (scientific) applications, but they target big systems and do not model the hardware constraints of MPSoC. Givy is a runtime system to execute dynamic task graphs on MPSoC. It has a focus on supporting irregular applications, and uses data-flow semantics to coordinate dynamic task scheduling and data transfer. To simplify the programmer's view of memory, both runtime and program data objects live in a GAS. To avoid address collisions when objects are dynamically allocated, and to maintain the consistency of these addresses across explicit data transfers and virtual memory remapping, a GAS-aware memory allocator is required. The allocator proposed in this work has the following properties: (1) it is free of inter-node synchronizations; (2) it is well suited for small memory systems; (3) its performances match that of existing state-of-the-art allocators.

6.2. On Characterizing the Data Access Complexity (IO) of Programs and Using it for Architectural Design Exploration

Participants: Venmugil Elango [OSU], Naser Sedaghati [OSU], Fabrice Rastello, Louis-Noël Pouchet [UCLA], J. Ramanujam [LSU], Radu Teodorescu [OSU], P. Sadayappan [OSU].

Technology trends will cause data movement to account for the majority of energy expenditure and execution time on emerging computers. Therefore, computational complexity will no longer be a sufficient metric for comparing algorithms, and a fundamental characterization of data access complexity will be increasingly important. The problem of developing lower bounds for data access complexity has been modeled using the formalism of Hong & Kung's red/blue pebble game for computational directed acyclic graphs (CDAGs). However, previously developed approaches to lower bounds analysis for the red/blue pebble game are very limited in effectiveness when applied to CDAGs of real programs, with computations comprised of multiple sub-computations with differing DAG structure. We address this problem by developing an approach for effectively composing lower bounds based on graph decomposition. We also develop a static analysis algorithm to derive the asymptotic data-access lower bounds of programs, as a function of the problem size and cache size.

The roofline model is a popular approach to "bounds and bottleneck" performance analysis. It focuses on the limits to performance of processors because of limited bandwidth to off-chip memory. It models upper bounds on performance as a function of operational intensity, the ratio of computational operations per byte of data moved from/to memory. While operational intensity can be directly measured for a specific implementation of an algorithm on a particular target platform, it is of interest to obtain broader insights on bottlenecks, where various semantically equivalent implementations of an algorithm are considered, along with analysis for variations in architectural parameters. This is currently very cumbersome and requires performance modeling and analysis of many variants.

We alleviate this problem by using the roofline model in conjunction with upper bounds on the operational intensity of computations as a function of cache capacity, derived using lower bounds on data movement. This enables bottleneck analysis that holds across all dependence-preserving semantically equivalent implementations of an algorithm. We demonstrate the utility of the approach in assessing fundamental limits to performance and energy efficiency for several benchmark algorithms across a design space of architectural variations.

This work is the fruit of the collaboration 8.4 with OSU. The first contribution (static analysis for lower bound) will be presented at ACM POPL'15 [10]. The second contribution (architectural exploration) is to be published at ACM TACO'15 [3].

6.3. A Tiling Perspective for Register Optimization

Participants: Duco Van Amstel, Lukasz Domagala, P. Sadayappan [OSU], Fabrice Rastello.

Register allocation is a much studied problem. A particularly important context for optimizing register allocation is within loops, since a significant fraction of the execution time of programs is often inside loop code. A variety of algorithms have been proposed in the past for register allocation, but the complexity of the problem has resulted in a decoupling of several important aspects, including loop unrolling, register promotion, and instruction reordering.

In this work, we develop an approach to register allocation and promotion in a unified optimization framework that simultaneously considers the impact of loop unrolling and instruction scheduling. This is done via a novel instruction tiling approach where instructions within a loop are represented along one dimension and innermost loop iterations along the other dimension. By exploiting the regularity along the loop dimension, and imposing essential dependence based constraints on intra-tile execution order, the problem of optimizing register pressure is cast in a constraint programming formalism. Experimental results are provided from thousands of innermost loops extracted from the SPEC benchmarks, demonstrating improvements over the current state-of-the-art.

This work is the fruit of both the collaboration 8.4 with OSU and with Kalray 7.1 7.2 .

6.4. Hybrid Data Dependence Analysis for Loop Transformations

Participants: Diogo Nunes Sampaio, Alain Ketterlin, Fabrice Rastello, Fernando Pereira, Alexandros Labrineas, Péricles Alves, Fabian Gruber.

Loop optimizations such as tiling, vectorization, or parallel task extraction are extremely important to achieve high performance. All such transformations rely on accurate memory dependence information to assess their validity. There are many practical situations, though, where dependence analysis fails to provide precise enough information. In this common scenario, the compiler will conservatively choose not to do any transformation. This happens in particular with low-level IRs (which are more and more common to address performance portability), but also in legacy code with pointers (e.g. C), linearized arrays, etc.

This work addresses the important problem of may-dependence disambiguation through the angle of a combination of static and dynamic analyses (sometimes called a hybrid analysis), similarly to what is already implemented in mainstream compilers, such as GCC, for auto-vectorization. This technique consists of adding a run-time test to disambiguate may-dependencies which static dependence analysis was not able to rule out. We propose two contributions to address this important problem.

The first approach proposes hybrid may-alias disambiguation. It combines two approaches: one that statically computes a symbolic expression of the interval of memory values a pointer may point to and uses dynamic overlap tests on these intervals to prove non-aliasing for each pair of pointers; another that hooks the memory allocator to find the base-pointer of a pointer and thus determine dynamically if a pointer pair belongs to two different allocations (and is thus disjoint) or not. We have applied these ideas on Polly-LLVM, a loop optimizer built on top of the LLVM compilation infrastructure. Our experiments indicate that our method is precise, effective and useful: we can disambiguate every pair of pointer in the loop intensive Polybench benchmark suite. The result of this precision is code quality: the binaries that we generate are 10% faster than those that Polly-LLVM produces without our optimization, at the -O3 optimization level of LLVM.

The second technique extends the non-overlapping intervals approach to may-dependence disambiguation. For this purpose, a powerful quantifier elimination scheme on multivariate-polynomials over integers has been developed. The quality of the presented scheme is important to make this approach realistic. In particular it must be precise (the integer aspect makes this problem very challenging), so that the test succeeds in practical cases, and must lead to negligible overhead. We evaluate preciseness and overhead on a set of 30+ benchmarks using complex loop transformations including loop fusion, skewing, and tiling.

This work is the fruit of the collaboration with UFMG 8.4 , Kalray 7.1 7.2 , STMicroelectronics 7.2 , and with EPI CAMUS in the context of IPL Multicore 8.2 . The first contribution has been presented at ACM OOPSLA'15 [19]. The second has been submitted to PLDI'16.

6.5. Power Efficiency and Computing Performance

Participants: Emilio Francesquini [UNICAMP, Campinas, Brazil], Edson Luiz Padoin [PhD: UFRGS and UNIJUI, Brazil], Marcio Castro [UFSC, Florianapolis, Brazil], Pedro Penna [PUC Minas, Belo Horizonte, Brazil], Henrique Cota de Freitas [PUC Minas, Belo Horizonte, Brazil], Fabrice Dupros [BRGM, Orléans, France], Philippe Navaux [UFRGS, Porto Alegre, Brazil], Jean François Méhaut.

Until the last decade, performance of HPC architectures has been almost exclusively quantified by their processing power. However, energy efficiency is being recently considered as important as raw performance and has become a critical aspect to the development of scalable systems. These strict energy constraints guided the development of a new class of so-called light-weight manycore processors. This study evaluates the computing and energy performance of two well-known irregular NP-hard problems – the Traveling-Salesman Problem (TSP) and K-Means clustering – and a numerical seismic wave propagation simulation kernel – Ondes3D – on multicore, NUMA, and manycore platforms. First, we concentrate on the nontrivial task of adapting these applications to a manycore, specifically the Kalray/MPPA-256 manycore processor. Then, we analyze their performance and energy consumption on those different machines. Our results show that applications able to fully use the resources of a manycore can have better performance and may consume from $3.8 \times$ to $13 \times$ less energy when compared to low-power and general-purpose multicore processors, respectively.

This work is the fruit of collaborations with Brazil and several universities (UFRGS, UFSC, UNICAMP, PUC Minas, USP). This work has been published in the journal of parallel and distributed computing [6] and in the journal of IET Computers and Digital Techniques [7]. This work was also part of several international projects (LICIA, CNPq/Inria HOSCAR project, Exase).

Emilio Francesquini and Marcio Castro are also former PhD students of University Grenoble Alpes (UGA) and the LIG Laboratory.

6.6. Modeling and Simulating of Dynamic Task-Based Runtime Systems

Participants: Luka Stanisic [PhD, Inria, Mescal], Samuel Thibault [Univ. Bordeaux, Inria, Storm], Brice Videau, Arnaud Legrand [CNRS, Inria, Mescal], Jean François Méhaut.

Multi-core architectures comprising several GPUs have become mainstream in the field of High-Performance Computing. However, obtaining the maximum performance of such heterogeneous machines is challenging as it requires to carefully offload computations and manage data movements between the different processing units. The most promising and successful approaches so far build on task-based runtimes that abstract the machine and rely on opportunistic scheduling algorithms. As a consequence, the problem gets shifted to choosing the task granularity, task graph structure, and optimizing the scheduling strategies. Trying different combinations of these different alternatives is also itself a challenge. Indeed, getting accurate measurements requires reserving the target system for the whole duration of experiments. Furthermore, observations are limited to the few available systems at hand and may be difficult to generalize. In this work, we show how we crafted a coarse-grain hybrid simulation/emulation of StarPU, a dynamic runtime for hybrid architectures, over SimGrid, a versatile simulator for distributed systems. This approach allows to obtain performance predictions of classical dense linear algebra kernels accurate within a few percents and in a matter of seconds, which allows both runtime and application designers to quickly decide which optimization to enable or whether it is worth

investing in higher-end GPUs or not. Additionally, it allows to conduct robust and extensive scheduling studies in a controlled environment whose characteristics are very close to real platforms while having reproducible behavior.

This work is part of the Luka Stanisic's thesis. Luka Stanisic was coadvised by Arnaud Legrand, Brice Videau and Jean-François Méhaut. This thesis was defended in November 2015. Luka Stanisic currently holds a postdoc position at Inria Bordeaux in the Storm and HiePacs teams. This work was published in the CCPE journal [9].

6.7. Fast and Accurate Simulation of Multithreaded Sparse Linear Algebra Solvers

Participants: Luka Stanisic [PhD, Inria, Mescal], Arnaud Legrand [CNRS, Inria, Mescal], Emmanuel Agullo [Inria, HiePacs], Alfredo Buttari [CNRS, IRIT, Toulouse], Florent Lopez [CNRS, IRIT, Toulouse], Brice Videau.

The ever growing complexity and scale of parallel architectures imposes to rewrite classical monolithic HPC scientific applications and libraries as their portability and performance optimization only comes at a prohibitive cost. There is thus a recent and general trend in using instead a modular approach where numerical algorithms are written at a high level independently of the hardware architecture as Directed Acyclic Graphs (DAG) of tasks. A task-based runtime system then dynamically schedules the resulting DAG on the different computing resources, automatically taking care of data movement and taking into account the possible speed heterogeneity and variability. Evaluating the performance of such complex and dynamic systems is extremely challenging especially for irregular codes. In this article, we explain how we crafted a faithful simulation, both in terms of performance and memory usage, of the behavior of `qr_mumps`, a fully-featured sparse linear algebra library, on multi-core architectures. In our approach, the target high-end machines are calibrated only once to derive sound performance models. These models can then be used at will to quickly predict and study in a reproducible way the performance of such irregular and resource-demanding applications using solely a commodity laptop.

This work is part of the Luka Stanisic's thesis. Luka Stanisic was coadvised by Arnaud Legrand, Brice Videau and Jean-François Méhaut. This thesis was defended in November 2015. Luka Stanisic currently holds a postdoc position at Inria Bordeaux in the Storm and HiePacs teams. This work was published in the ICPADS'2015 conference [18].

6.8. OpenMP Loop Scheduling

Participants: Pedro Penna [Master, PUC Minas, UFSC], Marcio Castro [Professor, UFSC], Henrique Cota de Freitas [Professor, PUC Minas], Francois Broquedis, Jean François Méhaut.

In High Performance Computing, the application's workload must be well balanced among the threads to achieve better performance. In this work, we propose a methodology that enables the design and exploration of new loop scheduling strategies. In this methodology, a simulator is used to evaluate the most relevant existing scheduling strategies, and a genetic algorithm is employed to explore the solution space of the problem itself. The proposed methodology allowed us to design a new loop scheduling strategy, which showed to be up to 32.3x better than the existing policies in terms of load balance.

6.9. BOAST: a Metaprogramming framework for computing kernels

Participants: Brice Videau [Postdoc CNRS, Mont-Blanc], Kevin Pouget [UJF, Nano2017], Luigi Genovese [Researcher, CEA INAC], Thierry Deutsch [Researcher, CEA INAC], Anthony Leonard [CNRS, Polytech Grenoble, Internship, from May 2015 until Aug 2015], Frederic Desprez, Jean François Méhaut.

Porting and tuning HPC applications to new platforms is tedious and costly in terms of human resources. Nonetheless, it is a very important aspect of the Mont-Blanc project. Indeed, for the Mont-Blanc project, more than ten applications were selected to be ported and optimized for the prototype platform.

Unfortunately, portability efforts are often lost when migrating to a new architecture. Worse, code may lose maintainability because several versions of some functionalities coexist, usually with a lot of duplication. Thus productivity of porting and tuning efforts is low as a huge fraction of those developments are never used after the platform they were intended for is decommissioned. Genericity of HPC codes is often limited. One of the reason is that producing generic code in Fortran 90/95 is difficult as the language does not really fit for it. Sometimes, adding genericity degrades performance as optimization opportunities that come from over-specification are lost. Functionality of HPC codes is tied to the previous point. Without genericity, adding new functionalities can be quite costly.

BOAST is a metaprogramming framework to produce portable and efficient computing kernels for HPC application. BOAST offers an embedded domain specific language to describe the kernels and their possible optimization. BOAST also supplies a complete runtime to compile, run, benchmark, and check the validity of the generated kernels. BOAST is being used in two flagship HPC applications BigDFT and SPECfem3D, to improve performance portability of those codes.

BOAST is developed in the context of Mont-Blanc projects. It will be also used in the C2S@Exa IPL and the H2020/HPC4E project.

6.10. Performance comparison between Java and JNI for optimal implementation of computational micro-kernels

Participants: Nassim Halli [PhD student, CIFRE Asetla Nanographics], Henri-Pierre Charles [CEA LIST, CRI PILSI], Jean François Méhaut.

General purpose CPUs used in high performance computing (HPC) support a vector instruction set and an out-of-order engine dedicated to increase the instruction level parallelism. Hence, related optimizations are currently critical to improve the performance of applications requiring numerical computation. Moreover, the use of a Java run-time environment such as the HotSpot Java Virtual Machine (JVM) in high performance computing is a promising alternative. It benefits from its programming flexibility, productivity and the performance is ensured by the Just-In-Time (JIT) compiler. Though, the JIT compiler suffers from two main drawbacks. First, the JIT is a black box for developers. We have no control over the generated code nor any feedback from its optimization phases like vectorization. Secondly, the time constraint narrows down the degree of optimization compared to static compilers like GCC or LLVM. So, it is compelling to use statically compiled code since it benefits from additional optimization reducing performance bottlenecks. Java enables to call native code from dynamic libraries through the Java Native Interface (JNI). Nevertheless, JNI methods are not inlined and require an additional cost to be invoked compared to Java ones. Therefore, to benefit from better static optimization, this call overhead must be leveraged by the amount of computation performed at each JNI invocation. In this work we tackle this problem and we propose to do this analysis for a set of micro-kernels. Our goal is to select the most efficient implementation considering the amount of computation defined by the calling context. We also investigate the impact on performance of several different optimization schemes which are vectorization, out-of-order optimization, data alignment, method inlining and the use of native memory for JNI methods.

This work was presented in the ADAPT'2015 workshop. It's also part of the Nassim Halli's thesis.

6.11. Reducing trace size in multimedia applications endurance tests

Participants: Serge Emteu [PhD ST Microelectronics, LIG/Slide, CORSE], Miguel Santana [ST Microelectronics], Alexandre Termier [Prof. Univ. Rennes I, IRISA/Inria/Dream], René Quiniou [CR Inria, IRISA/Inria/Dream], Brice Videau [Postdoc CNRS, Inria/Corse], Jean François Méhaut.

The consumer electronics market is dominated by embedded systems due to their ever-increasing processing power and the large number of functionalities they offer. To provide such features, architectures of embedded systems have increased in complexity : they rely on several heterogeneous processing units, and allow concurrent tasks execution. This complexity degrades the programmability of embedded system architectures and makes application execution difficult to understand on such systems. The most used approach for analyzing application execution on embedded systems consists in capturing execution traces (event sequences, such as system call invocations or context switch, generated during application execution). This approach is used in application testing, debugging or profiling. However in some use cases, execution traces generated can be very large, up to several hundreds of gigabytes. For example endurance tests, which are tests consisting in tracing execution of an application on an embedded system during long periods, from several hours to several days. Current tools and methods for analyzing execution traces are not designed to handle such amounts of data.

We propose an approach for monitoring an application execution by analyzing traces on the fly in order to reduce the volume of recorded traces. Our approach is based on features of multimedia applications which contribute the most to the success of popular devices such as set-top boxes or smartphones. This approach consists in identifying automatically the suspicious periods of an application execution in order to record only the parts of traces which correspond to these periods. The proposed approach consists of two steps : a learning step which discovers regular behaviors of an application from its execution trace, and an anomaly detection step which identifies behaviors deviating from the regular ones.

The many experiments, performed on synthetic and real-life datasets, show that our approach reduces the trace size by an order of magnitude while maintaining a good performance in detecting suspicious behaviors.

This work was presented at the DATE conference in Grenoble. It was also part of the Serge Emteu's thesis with ST Microelectronics.

6.12. Data Mining Approach to Temporal Debugging of Embedded Streaming Applications

Participants: Oleg Iegorov [PhD ST Microelectronics, LIG/Slide, CORSE], Miguel Santana [ST Microelectronics], Alexandre Termier [Prof. Univ. Rennes I, IRISA/Inria/Dream], Vincent Leroy [Associate Professor UJF, LIG/Slide], Jean François Méhaut.

One of the greatest challenges in the embedded systems area is to empower software developers with tools that speed up the debugging of QoS properties in applications. Typical streaming applications, such as multimedia (audio/video) decoding, fulfill the QoS properties by respecting the realtime deadlines. A perfectly functional application, when missing these deadlines, may lead to cracks in the sound or perceptible artifacts in the image.

We start from the premise that most of the streaming applications that run on embedded systems can be expressed under a dataflow model of computation, where the application is represented as a directed graph of the data flowing through computational units called actors. It has been shown that in order to meet real-time constraints the actors should be scheduled in a periodic manner. We exploit this property to propose SATM – a novel approach based on data mining techniques that automatically analyzes execution traces of streaming applications, and discovers significant breaks in the periodicity of actors, as well as potential causes of these breaks. We show on a real use case that our debugging approach can uncover important defects and pinpoint their location to the application developer.

This work was presented at the EMSOFT conference in Amsterdam. It was also part of the Oleg Iegorov's thesis with ST Microelectronics.

6.13. Tiling Bitwise Computations Using Look-up Instructions

Participants: Florent Bouchez - Tichadou, Cyril Six [Inria, Internship, from Feb 2015 until Jun 2015].

The BWLU is an instruction of a Very Long Instruction Word processor (VLIW) that performs a series of bit-independent computations in only one step through the use of a “look-up table” (LUT). The Bit-Wise Look-Up table instruction (BWLUT) takes as input four registers as well as a 32-bit integer (the “table”), and is able to output two bit-independent computations based on the input registers into two output registers.

The goal is to make the best use possible of this instruction by replacing during compilation as much as possible groups of bitwise computation using BWLUTs so as to reduce the number of instructions required to perform a computation. The problem is represented by a data-flow graph representing a computation, and the goal is use BWLUTs as tiles to “match” groups of bitwise instruction.

We proved the problem NP-complete for a general data-flow graph, so it is not practical to try to find the optimal solution.

It is easy to devise a greedy algorithm that will produce a solution, but we wanted a way to check whether the solutions found were far from the optimal. An optimal algorithm is of course exponential in the size of the input graph, however, we devised a complete space exploration algorithm based on dynamic programming that manages to find the optimal solution for data-flow graphs with small width or height.

CRYPT Team (section vide)

DECENTRALISE Team

6. New Results

6.1. Asynchronous Messaging

There are now a variety of end-to-end encrypted messaging platforms targeted at personal correspondences. Amongst these, only Pond and Ricochet provide meaningful resistance to traffic analysis by explicitly protecting the message metadata, although several can optionally operate over Tor to protect the user's location. Ricochet's design around Tor hidden services does not permit offline operation. Pond depends upon a centralized server.

In addition, there are messengers designed for academic research, like Vuvuzela, Dissent, and DP5. These employ information theoretically secure channels like dining cryptographers networks (DC-nets) and private information retrieval schemes (PIR) because they admit extremely simple proofs of security. As DC-nets and PIR schemes scale quadratically, these messaging research projects are effectively limited to a fixed maximum number of users, so they cannot realistically provide an alternative to modern email.

Instead, we have begun exploring the prospects of using mid-latency store-and-forward mixnets for asynchronous messaging. In fact, these are the amongst oldest anonymity systems, dating back to David Chaum, but they were normally restricted to anonymous email projects. At present, we remain in the early design phase, but our design scales linearly while providing many interesting properties desired by modern messengers.

We obtain provable security by basing our system on the Sphinx mixnet packet format, which is provably secure in the universal composability framework [7]. At first blush, Sphinx appears to be an overly restrictive format, but the restrictions are worth obtaining this degree of provable security along with a mixnet's scalability. After consideration, we have devised methods for adding entropy, and optimizing the location of entropy in Sphinx packet headers, without the need to use a larger and slower elliptic curve.

In Sphinx, there is a facility for single-use reply blocks (SURBs), as in other mixnets initially designed for anonymous remailers whose forward and backward messages look alike. We can store an SURB in the packet header, which enters use when the packet passes a fixed cross-over node, thereby allowing both sender and receiver remain anonymous to one another. We can orchestrate the usage of SURBs, and an authentication scheme using tokens, to provide optimal messaging properties that:

- Protect the identities of senders and recipients from each other and mixnet nodes, including the mailbox servers,
- Protect the identities of recipient's mailbox servers from even their contact to prevent denial of services attack,
- All redundancy through the usage of multiple mailbox servers.

We shall employ the Axolotl ratchet for long-term forward secrecy in messages, like Pond and Signal do. We can slightly improve upon the Axolotl ratchet by judiciously introducing side key material into the ratchet state. These side keys could be symmetric keys that take a different route through the mixnet, or travel outside the mixnet, thereby allowing the ratchet state to evolve based upon multiple concurrent paths. Side keys could also employ post-quantum public key cryptography, thus providing forward-secrecy against future attackers equipped with quantum computers.

We have also found another forward-secure ratchet inspired by Axolotl that integrates well with the Sphinx packet format. We believe this allows mixnet messages to be protected by long-term ratchets and posses a modicum of protection even against attackers with quantum-computers. At best, long-term ratchets themselves are only pseudonymous, not actually anonymous, so using the integrated ratchets requires considerable care.

6.2. Efficient Privacy-Preserving Scalar Product

We have designed, implemented and evaluated two variants of new privacy-preserving scalar product protocols. The first variant is based on an original idea of Ioannidis et al. [8] and was refined by Amirbekyan et al. [6]. Our first design improves on this by supporting signed values. A second design uses discrete logarithms over Elliptic curves instead of a homomorphic cipher, resulting in a substantially more efficient computation as long as the final result is numerically small.

In both protocols, Alice learns the scalar product $\sum a_i b_i$ of her private vector \vec{a} with Bob's private vector \vec{b} . The protocol is privacy-preserving in that Alice cannot discern details about \vec{b} other than what she can learn from \vec{a} and the scalar product $\sum a_i b_i$, and Bob does not learn anything.

Table 1 summarizes our experimental results.

Table 1. Preliminary performance data for the SP algorithms, wall-clock time running on a single-core of an i7.

Length	RSA-2048	ECC-2 ²⁰	ECC-2 ²⁸
25	14 s	2 s	29 s
50	21 s	2 s	29 s
100	39 s	2 s	29 s
200	77 s	3 s	30 s
400	149 s	OOB	31 s
800	304 s	OOB	33 s
800	3846 kb	OOB	70 kb

6.3. GNS support for Tor

We have worked with the Tor community to understand how best to support integration of the GNU Name System with Tor via specialized Tor exit nodes. There are two components to this work:

At present, there are somewhat fragile configuration options to Tor that should allow Tor users to locate the specialized exit nodes, although a small patch to Tor itself would improve upon these.

There are security reasons why Tor should not interact with locally configured name resolution services. OnionNS created a method to make Tor use local services for some domain name lookups, but doing so is somewhat heavy [9]. If we're creating a GNS patch to Tor anyways, then we'll likely extend it to optimize this process.

DEDUCTEAM Team

7. New Results

7.1. Termination

In [15], Frédéric Blanqui showed how to extend the notion of reducibility introduced by Girard for proving the termination of β -reduction in the polymorphic λ -calculus, to prove the termination of various kinds of rewrite relations on λ -terms, including rewriting modulo some equational theory and rewriting with matching modulo $\beta\eta$, by using the notion of computability closure. This provides a powerful termination criterion for various higher-order rewriting frameworks, including Klop's Combinatory Reductions Systems with simple types and Nipkow's Higher-order Rewrite Systems.

In [16], Frédéric Blanqui, together with Jean-Pierre Jouannaud and Albert Rubio, introduced the computability path ordering (CPO), a recursive relation on terms obtained by lifting a precedence on function symbols. A first version, core CPO, is essentially obtained from the higher-order recursive path ordering (HORPO) by eliminating type checks from some recursive calls and by incorporating the treatment of bound variables as in the so-called computability closure. The well-foundedness proof shows that core CPO captures the essence of computability arguments à la Tait and Girard, therefore explaining its name. We further show that no more type check can be eliminated from its recursive calls without losing well-foundedness, but one for which we found no counterexample yet. Two extensions of core CPO are then introduced which allow one to consider: the first, higher-order inductive types; the second, a precedence in which some function symbols are smaller than application and abstraction.

Another extension of CPO, to dependently typed terms, has been developed by Jean-Pierre Jouannaud and Jianqi Li in [50].

Jean-Pierre Jouannaud and Albert Rubio showed in [51] how to modify recursive path orders for higher-order terms which, like CPO, include $\beta\eta$ -reductions, into orders that are compatible with $\beta\eta$ -conversion. The result is a powerful order for proving termination of higher-order rewrite rules based on higher-order pattern matching.

Gaëtan Gilbert and Olivier Hermant have introduced a constructive way to perform proof normalization through completeness proofs [23].

Frédéric Blanqui formalized Ramsey's proof of the (infinite) Ramsey's theorem [54] (see <http://color.inria.fr/>).

7.2. Confluence

Jean-Pierre Jouannaud, in collaboration with Jiaxiang Liu, has started a program in order to enable confluence proofs in $\lambda\Pi$ modulo, investigating several open confluence problems for non-terminating relations. In [27], together with Mizuhito Ogawa, they introduced the new class of layered rewrite systems, and showed that their confluence can be reduced to that of their critical pairs computed by using unification over infinite rational terms when they do not increase the layer-depth of terms. This shows why an old example of non-terminating, left non-linear, critical pair free rewrite system due to Klop was non-confluent: it indeed had a critical pair in infinite rational trees. In the same paper, they also give an example of a non-confluent, layer-depth increasing system which has no critical pairs, hence showing that layer-depth plays a key role.

7.3. Automated theorem proving

In [25], Guillaume Bury, Raphaël Cauderlier and Pierre Halmagrand presented the extension of the automated theorem prover Zenon to ML-style polymorphism.

In [20], Guillaume Bury, David Delahaye, Damien Doligez, Pierre Hamalgrand and Olivier Hermant introduced an encoding of the set theory of the B method using polymorphic types and deduction modulo, used for the automated verification of proof obligations in the framework of the BWare project.

In [24], Kailiang Ji designed a strategy to translate model-checking problems into proving the satisfiability of a set of first-order formulas. The focus is to give an encoding of temporal properties expressed in CTL as first-order formulas, by translating the logical equivalence between temporal operators into rewrite rules. In this way, proof-search algorithms designed for Deduction Modulo, such as Resolution Modulo or Tableaux Modulo, can be used to verify temporal properties of finite transition systems. This strategy is implemented in iProver Modulo, and the testing results show that Resolution Modulo can be considered as a new way to quickly determine whether a temporal property is violated or not in transition system models.

7.4. $\lambda\Pi$ modulo and Dedukti

Gaëtan Gilbert, supervised by Arnaud Spiwack, wrote a prototype of a principle unification and type inference mechanism for Dedukti, based on a monadic API. This prototype separates with an abstraction barrier a unifier kernel which implements correct unification primitives from the unification algorithm and heuristics. The unification algorithm is written in a style which closely mirrors a pen-and-paper deduction rule presentation.

Éric Uzena, supervised by David Delahaye and Arnaud Spiwack, wrote a prototype of an extension of Dedukti with associative and commutative symbols and rewriting modulo associativity and commutativity of these symbols.

7.5. Encodings into Dedukti and interoperability

Ali Assaf, Guillaume Burel, Raphaël Cauderlier, David Delahaye, Gilles Dowek, Catherine Dubois, Frédéric Gilbert, Pierre Hamalgrand, Olivier Hermant, and Ronan Saillard have written a synthetic paper on the Dedukti system and on the expression of theories in this system. This paper is submitted to publication.

Ali Assaf [32] proved that Cousineau and Dowek's embedding of functional pure type systems [41] is conservative with respect to the original systems, using a new notion of reducibility called relative normalization. Together with Cousineau and Dowek's original result on the preservation of typing, this result justifies the use of the $\lambda\Pi$ -calculus modulo as a logical framework.

Ali Assaf's translation of the calculus of inductive constructions to the $\lambda\Pi$ -calculus modulo, which was presented at the TYPES conference in 2014, has been published in the postproceedings of TYPES 2014 [39]. This translation, which is based on the translation of pure type systems by Cousineau and Dowek [41], is implemented in the automated translation tool Coqine.

Ali Assaf and Guillaume Burel presented their translation of HOL to Dedukti at the PxTP 2015 workshop [18]. This translation, which is based on the translation of pure type systems by Cousineau and Dowek [41], is implemented in the automated translation tool Holide.

Raphaël Cauderlier and Catherine Dubois' translation of object calculus and subtyping to Dedukti, which was presented at the TYPES conference in 2014, has been published in the post-proceedings of TYPES 2014 [34].

In [26], Raphaël Cauderlier and Pierre Halmagrand presented a shallow embedding into Dedukti of proofs produced by ZenonModulo, an extension of the tableau-based first-order theorem prover Zenon to deduction modulo and typing.

In [33], Ali Assaf and Raphaël Cauderlier have combined simple developments written in Coq and HOL using Dedukti and the existing translation tools Coqine and Holide. This work is a first step towards using Dedukti as a framework for proof interoperability.

7.6. Proof theory

Guillaume Burel, Gilles Dowek and Ying Jiang have introduced a general framework to prove the decidability of reachability and provability problems. This framework uses an analogy between the objects recognized by an automaton and cut-free proofs. Various aspects of this work have been published at FroCoS [19], LPAR [21], and another paper is in preparation.

Gilles Dowek's paper on the definition of the classical connectives and quantifiers has been published [30].

Arnaud Spiwack gave a predicative shallow embedding of a weak version of system U^- in dependent type theory, for Hurkens's paradox to hold. He also showed that a variety of incarnations of Hurkens's paradox are straightforward instantiations of this encoding, greatly simplifying existing proofs.

Arnaud Spiwack developed a topos-theoretic methodology to reason equationally on circuit languages. Results that hold for combinational circuits are lifted to sequential circuits thanks to a transfer principle. This approach allows, in particular, to simplify reasoning about more complex temporal gates than the unit delay. These results aim at enriching the compiler of the Faust audio signal processing programming language, which features such complex temporal gates.

For the sake of reliability, the kernels of Interactive Theorem Provers (ITPs) are kept relatively small in general. On top of the kernel, additional symbols and inference rules are defined. Some dependency analysis of symbols of HOL Light indicates that the depth of dependency could be reduced by introducing a few more symbols to the kernel. Shuai Wang showed that extending the kernel of HOL Light is a successful attempt to reduce proof size and speed up proof-checking. More specifically, symbols and inference rules of universal quantification and implication were added to the kernel. This approach has been proved to give equivalent proof-checking results with the size of the proof files reduced to 24% on average and a speedup of 38% for proof-checking overall.

7.7. Computation models

Pablo Arrighi and Gilles Dowek have studied the expression of mechanic motions in cellular automata. Part of this work has been published in TPNC [17] and another paper is in preparation.

Arnaud Spiwack developed a variant of Turing machine where the tape is replaced by an unlabeled tree. The additional structure makes combining machines much easier, making it tractable to give explicit descriptions of rather complex machines. The cost model of these machines models that of purely functional programming languages, making it possible to compare mathematically the complexity of imperative algorithms and of purely functional algorithms.

DICE Team

6. New Results

6.1. The economy of intermediation and the anthropocene

Better understanding the economy, in a broad sense, of intermediation as it is performed by online platforms, is one of the major goals of the team. The paper [12] published in 1024, introduces the topics of algorithmic intermediation and its social impact to a large audience.

Two contemporary revolutions are shaking the world. On one side, the digital revolution, which seems to introduce to a new economic era, allowing more sharing, and according to some authors the end of capitalism. On the other hand, the challenges of the preservation of our planet, and the limitation of resources that we are now facing. Clearly, there is an expectation that digital means will help face the challenges of the planet. In [14], we go one step further and analyse the possible relationship between the two phenomena, by drawing comparisons with biology where stress on resources can lead to a horizontalisation of the species, much like what happens with digital technologies and intermediation platforms.

This later work is made in the framework of the study of the anthropocene, for which we are involved in the organisation of a workshop in the framework program of the HKW in Berlin on the technosphere

- URL: http://www.hkw.de/en/programm/projekte/2015/curriculum_campus_technosphere/campus_the_technosphere_issue.php

6.2. Geopolitics of intermediation platforms

Our study of the geopolitics of intermediation aims at grasping the balance of power between platforms, as well as between states - in their relation to platforms - and between platforms and states. We have designed coarse metrics [1] which capture the importance of a platform and the importance of a country in the digital landscape.

Our study focuses on the top 25 websites in a hundred countries. We emphasize the weight of intermediation platforms on the web. We also underline the imbalance between two digital powers - the United States and China - and the rest of the world. Indeed, most platforms belong to these two countries. We have extended our study to a deep analysis of the Asian case [8]. We develop our analysis in an interdisciplinary context as we collaborate with cartographers and economists. Two outcomes of our work are especially notable:

- We produce a set of maps and data visualisations to illustrate the intermediation economy [11].
- We highlight the determinants of the imbalance in the intermediation landscape. National policies and incentives are of primer importance. The digital landscapes of Korea and Taiwan for instance, show that countries can still play a main role in their domestic web [8].

6.3. Public administration and intermediation platforms

Building on the success of platforms such as Uber and the analyses of their externalities, we study the potential role of platforms in public administration. Indeed, cities such as Boston exhibit the interest of a collaboration between administrations and platforms in city planning and maintenance. We also address the role of platforms at a wider level as we study cases such as the settlement of the right to be forgotten in Europe. Our work benefits from the collaboration with administrations, such as Lyon metropole and social scientists. In particular, we have designed three possible scenarios of collaboration between platforms and institutions:

- Coexistence: platforms and institutions ignore themselves;
- Conflict: the services developed by platforms conflict with existing policies and institutional practices;
- Partnership: platforms and institutions partner around the development and promotion of services.

A working group has been established on digital sovereignty with CLTC, Center for Long Term Cybersecurity at UC Berkeley, Chaire Castex at Ecole Militaire, and Dice. This working group aims at getting a better understanding of the concept as well as the discrepancy of perception on both sides of the Atlantic. A first seminar was organised in Les Houches in december 2015.

This is work in progress with both academic and public administration actors.

6.4. Architecture design for intermediation platforms

Dice team designs software architectures for intermediation platforms. C3PO and BitBallot targets spontaneous and ephemeral social networks whereas Jumplyn focuses on pure central based system. All these architectures share a common JavaScript layout both at the client and the server sides. In the research context we validate state-of-the art technologies promoted by web leaders such as Google AngularJS, Facebook ReactJS and many others such as Netflix, Wallmart, or the Linux foundation for node.js. The Web environment raises many big issues since all equipments are basically connected to the Internet and the balance between end-user equipment cost and processing power is still a work in progress. Our main research track in such context is to find proper software toolkits hiding Web complexity. We mainly focus on time jitter, cornerstone of Web development, since it implies both end-user and network TCP indecisions. Due to this jitter combination the Web programming model has mutated toward the promises paradigm. It is a complex event based development model provided without external API help. It handles future execution whether successful or not, in a time jittered context. AngularJS, ReactJS, CoffeeScript, NodeJS, MongoDB, ElasticSearch are all time jitter compliant technologies designed for the Web constrains and revolutionising the way we build intermediation platforms.

C3PO explores network transport laziness with the use of a DTN that imposes a larger jitter than classical TCP/UDP. We build a JavaScript mockup [5] that uses a Java based DTN that stores, carries and forwards message from source to destination. C3PO is a software framework extending AngularJS through plugins, without central server, even during deployment phases. We use the dynamic nature of JavaScript to build application on the fly from network messages containing the application description. Our C3PO architecture enables us to build ephemeral and spontaneous social network, on demand and in a matter of days.

Our joint work with Worldline explores the promises paradigm model to enable automation extraction of independent micro-service. These micro-services called *fluxion* [9], from the contraction of flow and functions, may be dynamically and transparently moved over a cluster of servers. Our novelty resides in the fact that the original code is not redesigned for the cluster architecture. *Fluxion* are extracted from the initial code, and an equivalence is maintained between the initially *promissified* code and the *fluxionized* one. Code has two facets, a promise one, used to express software services and a *fluxion* one, used to express software bottlenecks.

Eventually our work with Jumplyn explores complex centralised social network. We want to design a software system to later support our technical research hot topics. The target theme is a software platform that helps students handle their projects. University depends more and more on external resources to teach students. Theses resources are both known by students and their teachers, and the pace and range of explored technologies leads to difficulties in teaching state-of-the-art subjects. The more dedicated a professor needs to be in his research activity, the more broad knowledge he has to teach. For instance 20 years ago one could cope software development teaching with one or two programming languages. Nowadays, a single code involves more then four programming languages to be fully understood. This technology spreading issue stands still in many teaching domains, since past technologies are still actives and future one are promising. We build Jumplyn to cope with this unbalanced game. To help student improving their project and avoid working with obsolete technologies, and to help teacher face the universal and inexpensive availability of knowledge. Jumplyn is a complex JavaScript development stack that collects resources for improving student work and providing services to help them from day to day activities. The current stack integrates the following technologies : MaterialDesign, AngularJS, CoffeeScript, NodeJs, MongoDb, ElasticSearch. Managing and developing software service above this stack is a complex research issue for a small sized development team. We do not have any publication on Jumplyn since our first goal is to build a support intermediation platform to

study classical issues such as recommendation or web crawling, scraping and indexation with our own sources of raw data.

DREAMPAL Project-Team

6. New Results

6.1. HoMade in 2015

6.1.1. Interruption support

In the last release of HoMade we introduced interruptions. Up to 7 interruptions are supported. The priority is static and each trap is associated to one of the 7 first VCs of the master, they are called trap1 .. trap7. Trap is par nature reflective. When a trap is raised the HoMade master reaches a no-preemptive kernel. Traps have no effect on the slaves, they can continue to work. At the end of trap execution, HoMade master resumes the sequential execution, trap codes should be clean and should reconstitute the stack as it was when they began. A WAIT instruction and a long IP cannot be interrupted. An example of interrupts is provided in the reconfiguration part later.

6.1.2. New assembly language

HoMade waits for two binary codes: one for the master and one for the slaves. These two codes are loaded via the UART port and triggers a global reset of all the softcores after. Binary codes are a sequence of 16 bits words finishing by a long word filled with 4 NULL. Our post fixed macro assembler generates some binary codes from text files. This assembly language introduces some flow controls like if for repeat. It is also based on PC and VC definitions. Now the particular operator := generates reflective behaviors via WIM instructions. The syntax is so simple than everybody can understand a program. A full new syntax description is available with the assembler on the official HoMade web site : <https://sites.google.com/site/homadeguide/assembleur-homade-v6>. Here is the code for a mono HoMade to implement a reflective execution of Fibonacci suite. Switches values are put on the top on the stack to indicate the position in the list we want to process. Different input buttons affect the execution: • Button 0 changes to soft fibo execution using some library IPs. SWAP ROT DUP = - + are IPs to change the tops of the stack or to process dyadic integer operators. • Button 1 changes to hard execution using fibo vhdl long IP • Other buttons process the current fibo (hard or soft).

```

:IP fibo $AC54 ; // fibo hard IPcode 54
// XX = 1 YY = 1
program
  : read
    $1f // immediate hexa
    btnpush // IP reads buttons pushed
    switch // IP reads switches
  ;
  : fibo_soft // function declare
    1 1 rot
    3 -
    for
      dup rot +
    next
    swap
    drop
  ;
  VC fibo_dyn := fibo_soft // VC init soft
start
begin
read

```

```

swap dup
0 = // test button
if // reflective process
    fibo_dyn := fibo_soft
endif
1 =
if
    fibo_dyn := fibo
endif
fibo_dyn // call VC
7seg // IP to print result
$1f
btn // button to pause
7seg
again // infinite loop
endprogram

```

When the VC `fibo_dyn` is called, you call hard or soft Fibonacci version depending of the sequence of pushed button. The soft code is 7 time slower than the hard code. The extra cost due to reflective facility is 2 cycles by VC call.

6.1.3. Dynamic IP reconfiguration

Xilinx chips are offering capabilities to program some pre-reserved chip areas with different bitstreams and this during the execution itself. It is not instantaneous and even worse the reconfiguration time depends of the length of the bitstream (the size of the area). Do not abuse of partial reconfigurations! But for some applications where context evolves at a “human speed”, our design can benefit of this functionality to adapt the hardware to the current context. It is easy to introduce this notion in HoMade: just insert an IP! This IP has to manage the bitstream memory and the ICAP to load them in the predefined areas. We develop a such IP for the master, without broadcast of bitstream to the slaves for the moment. This IP reconfiguration only needs to know the bitstream address. Effectively for Xilinx, the data inside the bitstream are sufficient to achieve the reconfiguration. We introduced the new keyword ‘in the assembler in order to express IP reconfigurations. The declaration of reconfigurable IPs may also include the bitstream address. Now we can program dynamic partial reconfiguration of IPs using our dedicated IP that we developed. Furthermore we can couple the dynamic reconfiguration with the reflective notion. Here is a simple example with dynamic image filters. The filter processes 1 block of 3x3 pixels. The 9 pixels are stored on the 3 top of the stack by aggregation of 3 pixels per word. External actuators can change from one IP to the other. We used interrupts and traps to apply this migration.

```

program // bistream addresses between ( )
:IP IP_median $EC11 ($0);
:IP IP_Sobel $EC22 ($49E);
VC filter
: T1
    IP_median ^^
    filter := IP_median
;
: T2
    IP_moyenne ^^
    filter := IP_moyenne
    trap1 := T1 // interrupt level 1
    trap2 := T2 // Interrupt level 2
    : get3pix // must be defined &
;
start

```

```
begin
  $7D for
    get3Pix // 3x3 pixels on stack
    get3Pix
    get3Pix
    -rot swap
  $7D for
    filter // current IP
    get3Pix // next 3 pixels
    -rot
  next
next
again
endprogram
```

Concerning dynamic reconfiguration of IPs, we are testing a dedicated IP to manage directly the ICAP of Xilinx. The different bitstreams are stored in DDR3 and this IP finds the starting address from the stack. Of course this is a long IP. Some optimization to broadcast efficiently the same bitstream towards different slave reconfigurable areas are still a big challenge with Xilinx architecture.

6.1.4. IP fusion

To be free from EDA companies, we are deploying IP fusion strategies to manage the dynamic reconfiguration by ourselves. We obtain good results concerning the reconfiguration time, but for large and very different IPs, the fusion works like an aggregation of two IPs and the surface gain is insignificant.

6.1.5. Using hardware parallelism for reducing power consumption in video streaming applications

In the PhD thesis of Karim Ali we exploited using a flexible parallel hardware-based architecture in conjunction with frequency scaling as a technique for reducing power consumption in video streaming applications. In this work, we derived equations to ease the calculation for the level of parallelism and the maximum depth for the FIFOs used for clock domain crossing. Accordingly, a design space was formed including all the design alternatives for the application. The preferable design alternative is selected in aware of how much hardware it costs and what power reduction goal it can satisfy. We used Xilinx Zynq ZC706 evaluation board to implement two video streaming applications: Video downscaler (1:16) and AES encryption algorithm to verify our approach. The experimental results showed up to 19.6% power reduction for the video downscaler and up to 5.4% for the AES encryption. The architecture and experimental results were published in a paper entitled "Using hardware parallelism for reducing power consumption in video streaming applications" at the 10th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC) in Jun 2015, Bremen, Germany [12].

In collaboration with NAVYA, we started the first steps to implement a stereo vision algorithm over a parallel architecture using FPGA technologies. The algorithm is based on a local approach for calculating the disparity map using sum of absolute difference between the right and the left image. As a first step, we exploited the possible optimization levels that can be applied at the software level. After that by using high level synthesis tool (Vivado HLS from Xilinx) the code was written in C in a way that facilitates its conversion into HDL files. Optimization techniques were applied to reduce both the hardware resources and time required for processing one frame. This design was tested experimentally to show around 50% decrease in the time required for processing one frame if compared to the software one. Currently, we are in the step of exploring more techniques for hardware optimization and decreasing the processing time to meet the industrial requirements of our partner.

6.1.6. A scalable flexible and dynamic reconfigurable architecture for high performance embedded computing

In collaboration with Nalam Embedded Systems (NES) and in the framework of the CIFRE PhD of Venkatasubramanian Viswanathan, we proposed a scalable and customizable reconfigurable computing platform, with a parallel full-duplex switched communication network, and a software execution model to redefine the computation, communication and reconfiguration paradigms in high performance embedded systems. High Performance Embedded Computing (HPEC) applications are becoming highly sophisticated and resource consuming for three reasons. First, they should capture and process real-time data from several I/O sources in parallel. Second, they should adapt their functionalities according to the application or environment variations within given Size Weight and Power (SWaP) constraints. Third, since they process several parallel I/O sources, applications are often distributed on multiple computing nodes making them highly parallel. Due to the hardware parallelism and I/O bandwidth offered by Field Programmable Gate Arrays (FPGAs), application can be duplicated several times to process parallel I/Os, making Single Program Multiple Data (SPMD) the favorite execution model for designers implementing parallel architectures on FPGAs. Furthermore Dynamic Partial Reconfiguration (DPR) feature allows efficient reuse of limited hardware resources, making FPGA a highly attractive solution for such applications. The problem with current HPEC systems is that, they are usually built to meet the needs of a specific application, i.e., lacks flexibility to upgrade the system or reuse existing hardware resources. On the other hand, applications that run on such hardware architectures are constantly being upgraded. Thus there is a real need for flexible and scalable hardware architectures and parallel execution models in order to easily upgrade the system and reuse hardware resources within acceptable time bounds. Thus these applications face challenges such as obsolescence, hardware redesign cost, sequential and slow reconfiguration, and wastage of computing power.

Addressing the challenges described above, we propose an architecture that allows the customization of computing nodes (FPGAs), broadcast of data (I/O, bitstreams) and reconfiguration several or a subset of computing nodes in parallel. The software environment leverages the potential of the hardware switch, to provide support for the SPMD execution model. Finally, in order to demonstrate the benefits of our architecture, we have implemented a scalable distributed secure H.264 encoding application along with several avionic communication protocols for data and control transfers between the nodes. We have used a FMC based high-speed serial Front Panel Data Port (sFPDP) data acquisition protocol to capture, encode and encrypt RAW video streams. The system has been implemented on 3 different FPGAs, respecting the SPMD execution model. In addition, we have also implemented modular I/Os by swapping I/O protocols dynamically when required by the system. We have thus demonstrated a scalable and flexible architecture and a parallel runtime reconfiguration model in order to manage several parallel input video sources. These results represent a conceptual proof of a massively parallel dynamically reconfigurable next generation embedded computers [16] [15]. The PhD of Venkatasubramanian Viswanathan has been defended in the 12th of october 2015 .

6.2. Language-Parametric Formal Methods

The HoMade assembly language is still evolving. Thus, our research in formal methods for programming languages kept the language-parametric nature that we decided upon when we started the project. The techniques and tools developed here will be instantiated on the HoMade assembly when it stabilizes. Our results are also applicable to general programming languages in order to target a broader audience.

In 2015 we have consolidated the results obtained in previous years, by making them more generally available and publishing them in high-end venues.

6.2.1. Language Definitions as Rewrite Theories

In [9] we study the foundations of \mathbb{K} , a formal framework for defining operational semantics of programming languages. The \mathbb{K} -Maude compiler translates \mathbb{K} language definitions to Maude rewrite theories. The compiler enables program execution by using the Maude rewrite engine with the compiled definitions, and program analysis by using various Maude analysis tools. \mathbb{K} supports symbolic execution in Maude by means of an

automatic transformation of language definitions. The transformed definition is called the *symbolic extension* of the original definition. In this paper we investigate the theoretical relationship between \mathbb{K} language definitions and their Maude translations, between symbolic extensions of \mathbb{K} definitions and their Maude translations, and how the relationship between \mathbb{K} definitions and their symbolic extensions is reflected on their respective representations in Maude. In particular, the results show how analysis performed with Maude tools can be formally lifted up to the original language definitions.

6.2.2. A Generic Framework for Symbolic Execution: Theory and Applications

In [10], [17] we propose a language-independent symbolic execution framework. The approach is parameterised by a language definition, which consists of a signature for the language's syntax and execution infrastructure, a model interpreting the signature, and rewrite rules for the language's operational semantics. Then, symbolic execution amounts to computing symbolic paths using a *derivative* operation. We prove that the symbolic execution thus defined has the properties naturally expected from it, meaning that the feasible symbolic executions of a program and the concrete executions of the same program mutually simulate each other. We also show how a coinduction-based extension of symbolic execution can be used for the deductive verification of programs. We show how the proposed symbolic-execution approach, and the coinductive verification technique based on it, can be seamlessly implemented in language definition frameworks based on rewriting such as the \mathbb{K} framework. A prototype implementation of our approach has been developed in \mathbb{K} . We illustrate it on the symbolic analysis and deductive verification of nontrivial programs.

6.2.3. Symbolic Execution by Language Transformation

In [2] we propose a language-independent symbolic execution framework for languages endowed with a formal operational semantics based on term rewriting. Starting from a given definition of a language, a new language definition is generated, with the same syntax as the original one, but whose semantical rules are transformed in order to rewrite over logical formulas denoting possibly infinite sets of program states. Then, the symbolic execution of concrete programs is, by definition, the execution of the same programs with the symbolic semantics. We prove that the symbolic execution thus defined has the properties naturally expected from it (with respect to concrete program execution). A prototype implementation of our approach was developed in the K Framework. We demonstrate the tool's genericity by instantiating it on several languages, and illustrate it on the reachability analysis and model checking of several programs.

6.2.4. Program Equivalence by Circular Reasoning

In [7] we propose a logic and a deductive system for stating and automatically proving the equivalence of programs written in languages having a rewriting-based operational semantics. The chosen equivalence is parametric in a so-called observation relation, and it says that two programs satisfying the observation relation will inevitably be, in the future, in the observation relation again. This notion of equivalence generalises several well-known equivalences and is appropriate for deterministic (or, at least, for confluent) programs. The deductive system is circular in nature and is proved sound and weakly complete; together, these results say that, when it terminates, our system correctly solves the given program-equivalence problem. We show that our approach is suitable for proving equivalence for terminating and non-terminating programs as well as for concrete and symbolic programs. The latter are programs in which some statements or expressions are symbolic variables. By proving the equivalence between symbolic programs, one proves the equivalence of (infinitely) many concrete programs obtained by replacing the variables by concrete statements or expressions. The approach is illustrated by proving program equivalence in two languages from different programming paradigms. The examples in the paper, as well as other examples, can be checked using an online tool.

6.2.5. Verifying Reachability-Logic Properties on Rewriting-Logic Specifications

Rewriting Logic is a simply, flexible, and powerful framework for specifying and analysing concurrent systems. Reachability Logic is a recently introduced formalism, which is currently used for defining the operational semantics of programming languages and for stating properties about program executions. Reachability Logic has its roots in a wider-spectrum framework, namely, in Rewriting Logic Semantics. In the invited

paper [10] we show how Reachability Logic can be adapted for stating properties of transition systems described by Rewriting-Logic specifications. We propose a procedure for verifying Rewriting-Logic specifications against Reachability-Logic properties. We prove the soundness of the procedure and illustrate it by verifying a communication protocol specified in Maude.

6.2.6. A Theoretical Foundation for Programming Languages Aggregation

This work was published as [11]. Programming languages should be formally specified in order to reason about programs written in them. We show that, given two formally specified programming languages, it is possible to construct the formal semantics of an aggregated language, in which programs consist of pairs of programs from the initial languages. The construction is based on algebraic techniques and it can be used to reduce relational properties (such as equivalence of programs) to reachability properties (in the aggregated language).

6.3. The SCAC Model : a weakly-coupled execution model for MPSoC

Synchronous Communication Asynchronous Computation (SCAC) is an execution model that separates the execution of communication phases from those of computation in order to facilitate their overlapping, thus covering the data transfer time. To allow the simultaneous execution of these two phases, we propose an approach based on three levels : two globally-centralized/locally-distributed hierarchical control levels and a parallel computation level.

G-MPSoC [5] is a SCAC System-on-Chip implementation based on a grid of clusters of Hardware and Software Computation Elements with different size, performance, and complexity. It is composed of parametric IP-reused modules: processor, controller, accelerator, memory, interconnection network, etc. to build different architecture configurations. The generic structure of G-MPSoC facilitates its adaptation to the intensive signal processing applications requirements.

The communication phase in SCAC System-on-Chip should be as fast as possible to avoid compromising parallel computing, using small and low power consumption modules to facilitate the interconnection network extensibility in a scalable system. To meet these criteria and based on a module reuse methodology, we chose to integrate a reconfigurable SCAC-Net [14] interconnection network to communicate data in our system. The SCAC-Net network is composed of communication modules as the number of the nodes used by the system. Using generic parameters, the topology of SCAC-Net network can be easily configured according to the needed communication which give more flexibility to the system.

ESTASYS Team

6. New Results

6.1. Heterogeneous Systems

Participants: Axel Legay, Jean Quilbeuf.

This part concerns Tasks 1, 2 and 4 of the action. We characterize and formalize heterogeneous aspects of SoS and then we define efficient monitoring algorithms and representations for their requirements. We then combine the results with Statistical Model Checking (Task 5).

Systems of Systems (SoS) are very large scale systems with particular characteristics. SoS are not directly built from scratch by a single designer or a single team but are obtained as the composition of simpler systems. SoS have strong reliability and dependability requirements, as they aim to provide a service over a long running period. SoS may dynamically modify themselves by connecting to new systems, updating or disconnecting faulty ones, making it impossible to statically know the set of subsystems that are part of the SoS before runtime.

One of the main difficulty arising when developing SoS is the fact that subsystems may have been designed with a different goal in mind. In particular, some subsystems may have their own goal which differs from the global goal of the SoS. Furthermore, each subsystem may be developed in a particular computation model, making it difficult to find a common unifying semantics for the whole SoS. Finally, SoS may exhibit some emergent behaviors that are hardly predictable at design time.

One of the solutions to allow simulation of an SoS is to rely on a common interface for interconnecting the subsystems. The Functional Mockup Interface (FMI) standard is a natural candidate for such an interface. The different components of an SoS developed in different models of computation can be translated to Functional Mockup Units (FMU). Then a so-called master algorithm coordinates the FMUs composing the system. The execution of each FMU is either directly handled by the master algorithm or relies on an external tool for its execution.

Because the subsystems composing an SoS are of heterogeneous nature, it is difficult to find a common semantics model for the whole system. Furthermore, building such a transition system is not tractable due to the complexity of the system. Thus verification through traditional model checking is not possible for SoS. However, since the FMI/FMU framework enables simulation of such systems, the statistical model checking approach can be used.

The DANSE EU project aims to provide a complete tool chain from the modeling to the verification of SoS. At the higher level, the modeling is done in UPDM using the RHAPSODY tool. At the same level, the designer can express requirements over the model using some patterns written in GCSL. The UPDM model can then be translated into a FMI/FMU format that can be simulated by a dedicated tool, named DESYRE. Similarly, the GCSL requirements are transformed into BLTL formulas. Finally, the PLASMA statistical model checker has been integrated with the DESYRE tool chain in order to check the BLTL formulas based on the simulations provided by DESYRE.

6.1.1. Papers:

papier DANSE(en cours) Ensuring a correct behaviour of SoS has a significant social impact. Their complexity and inherent dynamicity pose a serious challenge to traditional design methodologies. We propose a methodology and a tool-chain supporting design and validation of SoSs. We integrate SMC with existing industrial practice, by addressing both methodological and technological issues. Our contribution is summarized as follows: (1) a methodology for continuous and scalable validation of SoS formal requirements; (2) a natural-language based formal specification language able to express complex SoS requirements; (3) adoption of widely used industry standards for simulation

and heterogeneous systems integration (FMI and UPDM); (4) development of a robust SMC tool-chain integrated with system design tools used in practice. We illustrate the application of our SMC tool-chain and the obtained results on an industrial case study from the DANSE project.

6.2. Statistical Model Checking

Participants: Axel Legay, Sean Sedwards, Jean Quilbeuf, Louis-Marie Traonouez, Chan Ngo, Cyrille Jegourel.

This section covers Tasks 4 and 5 of the action. It consists in developing Simulation based techniques and efficient statistical algorithms for SoS.

The use of test cases remains the default means of ensuring the correct behaviour of systems in industry, but this technique is limited by the need to hypothesise scenarios that cause interesting behaviour and the fact that a reasonable set of test cases is unlikely to cover all possible eventualities. Static analysis is more thorough and has been successful in debugging very large systems, but its ability to analyse complex dynamical properties is limited. In contrast, model checking is an exhaustive technique that verifies whether a system satisfies a dynamical temporal logic property under all possible scenarios. For nondeterministic and probabilistic systems, numerical model checking quantifies the probability that a system satisfies a property. It can also be used to quantify the expected cost or reward of sets of executions.

Numerical model checking gives precise, accurate and certain results by exhaustively exploring the state space of the model, however the exponential growth of the state space with system size (the ‘state explosion problem’) typically limits its applicability to “toy” systems. Symbolic model checking using efficient data structures can make certain very large models tractable. It may also be possible to construct simpler but behaviourally equivalent models using various symmetry reduction techniques, such as partial order reduction, bisimulation and lumping. If a new system is being constructed, it may be possible to guarantee the overall behaviour by verifying the behaviour of its subcomponents and limiting the way they interact. Despite these techniques, however, the size, unpredictability and heterogeneity of real systems usually make numerical techniques infeasible. Moreover, even if a system has been specified not to misbehave, it is nevertheless necessary to check that it meets its specification.

Simulation-based approaches are becoming increasingly tractable due to the availability of high performance parallel hardware and algorithms. In particular, statistical model checking (SMC) combines the simplicity of testing with the formality of numerical model checking. The core idea of SMC is to create multiple independent execution traces of a system and count how many satisfy a property specified in temporal logic. The proportion of satisfying traces is an estimate of the probability that the system satisfies the property. By thus modelling the executions of a system as a Bernoulli random variable, the absolute error of the estimate can be bounded using, for example, a confidence interval or a Chernoff bound. It is also possible to use efficient sequential hypothesis testing, to decide with specified statistical confidence whether the probability of a property is above or below a given threshold. Since SMC requires multiple independent simulations, it may be efficiently divided on parallel computer architectures, such as grids, clusters, clouds and general purpose computing on graphics processors (GPGPU).

Knowing a result with less than 100% confidence is often sufficient in real applications, since the confidence bounds may be made arbitrarily tight. Moreover, a swiftly achieved approximation may prevent a lot of wasted time during model design. For many complex systems, SMC offers the only feasible means of quantifying performance. Historically relevant SMC tools include APMC, YMER and VESTA. Well-established numerical model checkers, such as PRISM and UPPAAL, are now also including SMC engines. Dedicated SMC tools under active development include COSMOS and our own tool PLASMA. Recognising that SMC may be applied to any discrete event trace obtained by stochastic simulation, we have devised PLASMA-lab, a modular library of SMC algorithms that may be used to construct domain-specific SMC tools. PLASMA-lab has become the main vehicle of our ongoing development of SMC algorithms.

Statistical model checking (SMC) addresses the state explosion problem of numerical model checking by estimating quantitative properties using simulation. To advance the state of the art of SMC we address the ongoing challenges of rare events and nondeterminism. We also make novel use of SMC by applying it to motion planning in the context of assisted living. Rare events are often of critical importance and are challenging to SMC because they appear infrequently in simulations. Nondeterministic models are useful to model unspecified interactions, but simulation requires that nondeterminism is resolved.

We also applied SMC in the context of Systems of Systems (SoS). In the frame of the DANSE project, Plasma-Lab was used to verify SoS, and completely integrated with the DANSE tool-chain. We are currently working on verification of dynamic SoS, where systems can appear and disappear during execution. This work is done in collaboration with the ArchWare team from IRISA. We will interface Plasma-Lab with a simulator for the Pi-ADL language that enables simulation of dynamic systems.

Our group is devising cutting edge techniques for SMC. In particular, we are developing new algorithms for non-deterministic systems as well as for dynamic systems. Rare event systems are also addressed. Finally, we also devote a large amount of time to applying our technology to realistic case studies described in high-level languages such as Simulink or System C, or even a robot moving an elderly person in a commercial center.

6.2.1. Papers:

- [2] (J) People with impaired physical and mental ability often find it challenging to negotiate crowded or unfamiliar environments, leading to a vicious cycle of deteriorating mobility and sociability. To address this issue we present a novel motion planning algorithm that is able to intelligently deal with crowded areas, permanent or temporary anomalies in the environment (e.g., road blocks, wet floors) as well as hard and soft constraints (e.g., “keep a toilet within reach of 10 meters during the journey”, “always avoid stairs”). Constraints can be assigned a priority tailored on the user’s needs. The planner has been validated by means of simulations and experiments with elderly people within the context of the DALi FP7 EU project.
- [3] (J) Markov decision processes (MDP) are useful to model optimisation problems in concurrent systems. To verify MDPs with efficient Monte Carlo techniques requires that their nondeterminism be resolved by a scheduler. Recent work has introduced the elements of lightweight techniques to sample directly from scheduler space, but finding optimal schedulers by simple sampling may be inefficient. Here we describe “smart” sampling algorithms that can make substantial improvements in performance.
- [21] (C) Rare properties remain a challenge for statistical model checking (SMC) due to the quadratic scaling of variance with rarity. We address this with a variance reduction framework based on lightweight importance splitting observers. These expose the model-property automaton to allow the construction of score functions for high performance algorithms. The confidence intervals defined for importance splitting make it appealing for SMC, but optimising its performance in the standard way makes distribution inefficient. We show how it is possible to achieve equivalently good results in less time by distributing simpler algorithms. We first explore the challenges posed by importance splitting and present an algorithm optimised for distribution. We then define a specific bounded time logic that is compiled into memory-efficient observers to monitor executions. Finally, we demonstrate our framework on a number of challenging case studies.
- [23] (C) Exhaustive verification can quantify critical behaviour arising from concurrency in nondeterministic models. Rare events typically entail no additional challenge, but complex systems are generally untractable. Recent work on Markov decision processes allows the extremal probabilities of a property to be estimated using Monte Carlo techniques, offering the potential to handle much larger models. Here we present algorithms to estimate extremal rewards and consider the challenges posed by rarity. We find that rewards require a different interpretation of confidence and that reachability rewards require the introduction of an auxiliary hypothesis test. We show how importance sampling can significantly improve estimation when probabilities are low, but find it is not a panacea for rare schedulers.

- [36] (J; accepted) We propose a new SMC technique based on CUSUM, an algorithm originally used in signal processing, that detects probability change at runtime on a single execution of a system. The principle is to monitor the execution at regular time intervals, and to perform Monte Carlo checks over the samples of the execution. The results of these checks are used to compute the CUSUM ratio, whose variation allows to detect a change of the probability measure of the system. We demonstrate the algorithm to detect failures in a Simulink model of a temperature controller. Computing the exact time at which failures may happen is then useful to schedule maintenance operations.
- [42] (W) Many embedded and real-time systems have a inherent probabilistic behaviour (sensors data, unreliable hardware,...). In that context, it is crucial to evaluate system properties such as “the probability that a particular hardware fails”. Such properties can be evaluated by using probabilistic model checking. However, this technique fails on models representing realistic embedded and real-time systems because of the state space explosion. To overcome this problem, we propose a verification framework based on *Statistical Model Checking*. Our framework is able to evaluate probabilistic and temporal properties on large systems modelled in SystemC, a standard system-level modelling language. It is fully implemented as an extension of the Plasma-lab statistical model checker. We illustrate our approach on a multi-lift system case study.
- [27] (W) Stochastic Petri nets are commonly used for modeling distributed systems in order to study their performance and dependability. This report proposes a realization of stochastic Petri nets in SystemC for modeling large embedded control systems. Then statistical model checking is used to analyze the dependability of the constructed model. Our verification framework allows users to express a wide range of useful properties to be verified which is illustrated through a case study.
- [25] (C; accepted) Transaction-level modeling with SystemC has been very successful in describing the behavior of embedded systems by providing high-level executable models, in which many of them have an inherent probabilistic behavior, i.e., random data, unreliable components. It is crucial to evaluate the quantitative and qualitative analysis of the probability of the system properties. Such analysis can be conducted by constructing a formal model of the system and using probabilistic model checking. However, this method is infeasible for large and complex systems due to the state space explosion. In this work, we demonstrate the successful use of *Statistical Model Checking* to carry out such analysis directly from large SystemC models and allows designers to express a wide range of useful properties. This work is going to be presented at 17th IEEE High Assurance Systems Engineering Symposium in January, 2016.

6.3. Formal Models for Variability

Participants: Axel Legay, Rudolf Fahrenberg, Jin Hyun Kim.

This part of the report is more concerned with task 2. It studies variability aspects in the broad scope. As in the first year, we have decided to use the concept of product lines as a general framework to reason on the problematic.

The behaviour of a software system is often described in terms of its features, where each *feature* is a unit of functionality that adds value to the system. *Feature-oriented software development (FOSD)* is a software-development strategy that is based on feature decomposition and modularity. Features can be separate modules that are developed in isolation, allowing for parallel, incremental, or multi-vendor development of features. Feature orientation is particularly important in *software product lines*, where a family of related products is managed and evolved in terms of its features: a product line comprises a collection of mandatory and optional features, and individual products are derived by selecting among and integrating features from this feature set. A product line can be expressed as a single model, in which feature-specific behaviour is conditional on the presence of the feature in a product.

The downside of FOSD is that, although features are conceptualized, developed, managed, and evolved as separate concerns, they are not truly separate. They can interfere with each other, for example by trying to control the same variables, by issuing events that trigger other features, or by imposing conditions that suppress other features. Most of the early work on feature interactions focused on interactions that manifest themselves as logical inconsistencies, such as conflicting actions, nondeterminism, deadlock, invariant violation, or unsatisfiability. More recently, a more general definition of feature interaction has been introduced, in terms of a feature that is developed and verified to be correct in isolation but is found to behave differently when combined with other features, and it was shown how such *behaviour interactions* could be detected as a violation of bisimulation.

Another problem is that FTS models are monolithic models of full product lines. There is no means of modelling individual features and composing them into products or product-line models, or of specifying feature increments to an existing product-line model. As such, FTSs cannot be the mathematical basis for modelling technologies that support feature decomposition, composition, or incremental evolution of a product line.

6.3.1. Papers:

- [11] (C) Featured Transition Systems (FTSs) is a popular representation for software product lines: an entire product line is compactly represented as a single transition-machine model, in which feature-specific behaviour is guarded by feature expressions that are satisfied (or not) by the presence or absence of individual features. In previous work, FTS models were monolithic in the sense that the modeller had to construct the full FTS model of the product line in its entirety. To allow for modularity of FTS models, we propose here a language for extending an existing FTS model with new features. We demonstrate the language using a running example and present results about the language's expressivity, commutativity of feature extensions, feature interactions, and resolution of such interactions.
- [12] (C) We suggest a method for measuring the degree to which features interact in feature-oriented software development. To this end, we extend the notion of simulation between transition systems to a similarity measure and lift it to compute a behaviour interaction score in featured transition systems. We then develop an algorithm which can compute the degree of feature interactions in a featured transition system in an efficient manner.

6.4. Privacy and Security

Participants: Axel Legay, Fabrizio Biondi, Jean Quilbeuf, Thomas Given-Wilson, Sébastien Josse.

6.4.1. Information-Theoretical Quantification of Security Properties

This part of the work was not foreseen at the beginning of the action. It concerns security aspects, and more precisely quantifying privacy of data. This aspect is in fact central for SoS and all our algorithms developed for Tasks 4 and 5 should be adapted to solve a series of problems linked to privacy in interconnected object and dynamical environment. For now, we only studied the foundations.

Information theory provides a powerful quantitative approach to measuring security and privacy properties of systems. By measuring the *information leakage* of a system security properties can be quantified, validated, or falsified. When security concerns are non-binary, information theoretic measures can quantify exactly how much information is leaked. The knowledge of such informations is strategic in the developments of component-based systems.

The quantitative information-theoretical approach to security models the correlation between the secret information of the system and the output that the system produces. Such output can be observed by the attacker, and the attacker tries to infer the value of the secret by combining this information with its knowledge of the system.

Armed with the produced output and the source code of the system, the attacker tries to infer the value of the secret. The quantitative analysis we implement computes with arbitrary precision the number of bits of the secret that the attacker will expectedly infer. This expected number of bits is the information leakage of the system.

The quantitative approach generalizes the qualitative approach and thus provides superior analysis. In particular, a system respects non-interference if and only if its leakage is equal to zero. In practice very few systems respect non-interference, and for those who don't it is imperative to be able to distinguish between the ones leaking a very small amount of bits and the ones leaking a significant amount of bits, since only the latter are considered to pose a security vulnerability to the system.

Since black box security analyzes are immediately invalidated whenever an attacker gains information about the source code of the system, we assume that the attacker has a white box view of the system, meaning that it has access to the system's source code. This approach is also consistent with the fact that many security protocol implementations are in fact open source.

The scope of modern software projects is too large to be analyzed manually. For this reason we provide tools that can support the analyst and locate security vulnerabilities in large codebases and projects. We work with a variety of tools, including commercial software analysis tools being adapted with our techniques, and tools such as QUAIL developed here by our team.

We applied the leakage analysis provided by QUAIL to several case studies. Our case studies (voting protocol and smart grid coordination) have in common that a publicly disclosed information is computed from the secret of every participant in the model. In the voting example, the vote of a given voter is secret, but the number of votes for each candidate is public. Similarly, in the smart grid example, the consumption of one of the houses is secret, but the consumption of a whole quarter can be deduced. Qualitative analyses are either too restrictive or too permissive on these types of systems. For instance, non-interference will reject them as the public information depends on the secret. Declassification approaches will accept them, even if the number of voters or consumers is 2, in which case the secret can be deduced.

The development of better tools for quantitative security builds upon both theoretical developments in information theory, and development of the tools themselves. These often progress in parallel with each supporting the findings of the other, and increasing the demands and understanding upon each other.

6.4.1.1. Papers:

- [34] (C; submitted) Systems dealing with confidential data may leak some information by their observable outputs. Quantitative information flow analysis provides a method for quantifying the amount of such information leakage. To avoid the high computational cost of exhaustive search, statistical analysis has been studied to estimate information leakage by analyzing only a small but representative subset of the system's behavior. In this paper we propose a new compositional statistical analysis method for quantitative information flow that combines multiple statistical analyses with static trace analysis. We use partial knowledge of the system's source code or specification, therefore improving both quality and cost of the analysis. The new method can optimize the use of weighted statistical analysis by performing it on components of the system and appropriately adapting their weights. We show this approach combined with the precision of trace analysis produces better estimates and narrower confidence intervals than the state of the art.
- [38] (J) The quantification of information leakage provides a quantitative evaluation of the security of a system. We propose the usage of Markovian processes to model deterministic and probabilistic systems. By using a methodology generalizing the lattice of information approach we model refined attackers capable to observe the internal behavior of the system, and quantify the information leakage of such systems. We also use our method to obtain an algorithm for the computation of channel capacity from our Markovian models. Finally, we show how to use the method to analyze timed and non-timed attacks on the Onion Routing protocol.
- [40] (C) Quantitative security analysis evaluates and compares how effectively a system protects its secret data. We introduce QUAIL, the first tool able to perform an arbitrary-precision quantitative

analysis of the security of a system depending on private information. QUAIL builds a Markov Chain model of the system's behavior as observed by an attacker, and computes the correlation between the system's observable output and the behavior depending on the private information, obtaining the expected amount of bits of the secret that the attacker will infer by observing the system. QUAIL is able to evaluate the safety of randomized protocols depending on secret data, allowing to verify a security protocol's effectiveness. We experiment with a few examples and show that QUAIL's security analysis is more accurate and revealing than results of other tools.

- [41] (C) Quantitative security techniques have been proven effective to measure the security of systems against various types of attackers. However, such techniques are based on computing exponentially large channel matrices or Markov chains, making them impractical for large programs. We propose a different approach based on abstract trace analysis. By analyzing directly sets of execution traces of the program and computing security measures on the results, we are able to scale down the exponential cost of the problem. Also, we are able to apply statistical simulation techniques, allowing us to obtain significant results even without exploring the full space of traces. We have implemented the resulting algorithms in the QUAIL tool. We compare their effectiveness against the state of the art LeakWatch tool on two case studies: privacy of user consumption in smart grid systems and anonymity of voters in different voting schemes.
- [37] (C) In an election, it is imperative that the vote of the single voters remain anonymous and undisclosed. Alas, modern anonymity approaches acknowledge that there is an unavoidable leak of anonymity just by publishing data related to the secret, like the election's result. Information theory is applied to quantify this leak and ascertain that it remains below an acceptable threshold. We apply modern quantitative anonymity analysis techniques via the state-of-the-art QUAIL tool to the voting scenario. We consider different voting typologies and establish which are more effective in protecting the voter's privacy. We further demonstrate the effectiveness of the protocols in protecting the privacy of the single voters, deriving an important desirable property of protocols depending on composite secrets.
- [39] (C) In recent years, quantitative security techniques have been providing effective measures of the security of a system against an attacker. Such techniques usually assume that the system produces a finite amount of observations based on a finite amount of secret bits and terminates, and the attack is based on these observations. By modeling systems with Markov chains, we are able to measure the effectiveness of attacks on non-terminating systems. Such systems do not necessarily produce a finite amount of output and are not necessarily based on a finite amount of secret bits. We provide characterizations and algorithms to define meaningful measures of security for non-terminating systems, and to compute them when possible. We also study the bounded versions of the problems, and show examples of non-terminating programs and how their effectiveness in protecting their secret can be measured.

6.4.2. Equivocation-based Security Measures for Shared-Key Cryptosystems

Ensuring privacy and security of communication is a fundamental concern of cyber-physical systems and handled by encryption. Information-theoretic reasoning allows the modelling of security properties via unconditional security. That is, information-theoretic approaches formalise security properties that do not rely upon unproven computational hardness results, and are not vulnerable to advances in computing hardware, software or theory. For example, such unconditional security guarantees are not weakened by quantum computers, mem-computers, or new mathematical discoveries.

Traditionally the strongest measure of the security of a system is *perfect secrecy* as proposed by Shannon. However, this relies upon having a large key that is used only once. In practice a measure of the security of cryptosystems that does not meet this requirement is more useful. To this end we presented *max-equivocation*, a measure of the maximum achievable security given the keys available. Indeed max-equivocation not only formalizes the best possible security, but also generalizes perfect secrecy.

Max-equivocation holds even when inputs to the systems (i.e. keys and messages) are not uniform. This corresponds to many real world scenarios, and indeed we have shown that existing approaches are non-optimal as they do not consider these perturbations in the inputs. We provide necessary and sufficient conditions for achieving max-equivocation, formalizing exactly when it can be achieved in practice.

We further generalize to consider scenarios where message spaces are not complete, i.e. there are messages that are invalid and could never be produced. This allows reasoning over (and contrasting with) many prior approaches as well as formalizing their strengths and weaknesses under max-equivocation.

The most common attack against such cryptosystems is to consider when the attacker sees a single (encrypted) message and tries to guess the content. This can be measured by the *vulnerability* of the system, i.e. the probability that the attacker will guess correctly the message. We formalize a *relative vulnerability* for when the attacker makes this guess under incorrect assumptions about the messages. We formalize that the attacker can never improve their chances at guessing the message with incorrect assumptions.

Now we consider what information the attacker can gain by observing the cryptosystem. We show that the encryption function alone reveals information about the possible message distributions to the attacker. In the worse case scenario an encryption function may admit only a single message distribution. Thus the construction of the encryption function should consider this and (when possible) admit many solutions.

Further we consider what the attacker can learn by observing the communication of a cryptosystem. We show that the attacker can learn the probability distribution over the ciphertexts (encrypted messages), and combined with the information from the encryption function converge upon a distribution for the messages. Again if the encryption function admits one solution then the attacker learns the exact message distribution. We show that even when a single solution will not be found, the attacker still converges upon a message distribution that can only improve their attacks.

In addition to formalizing how these attacks work, and thus how to protect against them when constructing cryptosystems, we also consider not sharing the encryption function as a mechanism to avoid the attacker exploiting it. We formalize how to still communicate effectively in this scenario, and the advantages and disadvantages of this approach.

We present several algorithms to demonstrate the practicality of the techniques. The algorithms to achieve max-equivocation consider the message distribution and compute an encryption function that achieves close to max-equivocation. Since these algorithms are tailored for the message distributions, they out perform generic algorithms. We also present algorithms that are able to perform well without revealing the entire encryption function, and thus revealing less information to the attacker and hindering cryptanalysis.

Thus we show that unconditional security is not only more resistant to technology changes, but also can be formalised for many scenarios, and is achievable in practice.

6.4.2.1. Papers:

- [29] (C, submitted) Recent work has presented max-equivocation as a measure of the resistance of a cryptosystem to attacks when the attacker is aware of the encoder function and message distribution. Here we consider the vulnerability of a cryptosystem in the one-try attack scenario when the attacker has incomplete information about the encoder function and message distribution. We show that encoder functions alone yield information to the attacker, and combined with inferable information about the ciphertexts, information about the message distribution can be discovered. We show that the whole encoder function need not be fixed or shared a priori for an effective cryptosystem, and this can be exploited to increase the equivocation over an a priori shared encoder. Finally we present two algorithms that operate in these scenarios and achieve good equivocation results, ExPad that demonstrates the key concepts, and ShortPad that has less overhead than ExPad.
- [13], [28] (C; J, submitted) Preserving the privacy of private communication is a fundamental concern of computing addressed by encryption. Information-theoretic reasoning models unconditional security where the strength of the results is not moderated by computational hardness or unproven results. Perfect secrecy is often considered the ideal result for a cryptosystem, where knowledge of the ciphertext reveals no information about the message or key, however often this is impossible to

achieve in practice. An alternative measure is the equivocation, intuitively the average number of message/key pairs that could have produced a given ciphertext. We show a theoretical bound on equivocation called max-equivocation and show that this generalizes perfect secrecy when achievable, and provides an alternative measure when perfect secrecy is not. We derive bounds for max-equivocation, and show that max-equivocation is achieved when the entropy of the ciphertext is minimized. We consider encryption functions under this new perspective, and show that in general the theoretical best is unachievable, and that some popular approaches such as Latin squares or Quasigroups are also not optimal. We present some algorithms for generating encryption functions that are practical and achieve 90 - 95% of the theoretical best, improving with larger message spaces.

6.4.3. Malware Classification via Deobfuscation and Behavioral Fingerprinting

A fundamental problem to guarantee the security of systems is to be able to discriminate between legitimate processes and processes with malicious behavior. Malicious software, or malware, has to be identified and prevented from executing on the system, and its actions reverted by a disinfection process. To be able to recognize and disinfect malware it is necessary to be able to extract a behavioral fingerprint or signature from a binary file, and to construct a database of such signatures for comparison. The signatures in the database have to be classified according to the malware's family and category, allowing the correct disinfection method to be deployed.

Automatic extraction of behavioral signatures in the form of temporal logical graphs or control flow graphs is a recent but very effective technique, and malware developers have already adapted malware compilation chains to include techniques to hinder reverse engineering and thus prevent the extraction of such signatures. These obfuscation techniques include the addition of obfuscated conditional statements leading to dead code, control flow flattening based on complex function like cryptographic hash functions, and source code virtualization on an embedded interpreter.

Consequently, deobfuscation has to be developed along with fingerprinting techniques to be able to effectively extract malware signatures. We are pushing the state of the art in both subjects, advancing generalized and targeted deobfuscation and deploying them on an innovative virtualization and malware fingerprinting tool.

Mixed Boolean Arithmetic (MBA) obfuscation is an obfuscation technique developed by Cloakware Inc. and deployed in obfuscating compilation chains for both legitimate code and malware. We have deployed state-of-the-art SMT solvers to evaluate their effectiveness against MBA-obfuscated conditionals and ascertained their limited effectiveness. So we have developed an algebraic simplification technique targeting the algebraic structure of MBA obfuscation, and proved such technique to be extremely effective, being able to deobfuscate statements in orders of magnitude less time than the time required to obfuscate them in the first place.

While the algebraic simplification technique is very effective against MBA obfuscation, it is completely tailored to MBA obfuscation. We instead explore a completely general method based on dynamic program synthesis. Synthesis algorithms, like the ones based on Reed-Muller expansion techniques, interrogate the target (in this case the obfuscated conditional) multiple times considering it as a black-box oracle, and synthesize the function expressed by the target from the answers to such interrogation. We determined that synthesis is significantly more efficient than SMT solving in synthesizing the obfuscated function in a very compact form, and thus very promising as a generalized deobfuscation method.

While more targeted deobfuscation techniques are required to counteract control flow flattening and virtualization, the deobfuscation of conditional statements is an important step for malware fingerprinting. We plan to use our tool to classify a large database of malware, producing an extensive database of malware signatures representing multiple versions and families of malicious code. Malware mining and evolution techniques can be deployed on such database to construct different signatures for unknown variants of similar malware, thus improving the effectiveness of the detection process.

6.4.3.1. Papers:

- [30] (C, submitted) The obfuscation of conditional statements is a simple and efficient way to disturb the identification of the control flow graph of a program. Mixed Boolean arithmetics (MBA) techniques provide concrete ways to achieve this obfuscation of conditional statements. In this work, we

study the effectiveness of automated deobfuscation of MBA obfuscation, using algebraic, SMT-based and synthesis-based techniques. We experimentally ascertain the practical feasibility of MBA obfuscation. We study using SMT-based approaches with different state-of-the-art SMT solvers to counteract MBA obfuscation, and we show how the deobfuscation complexity can be greatly reduced by algebraic simplification. We also consider synthesis-based deobfuscation and find it to be more effective than SMT-based deobfuscation. We discuss complexity and limits of all methods, and conclude that MBA obfuscation is not effective enough to be considered a reliable method for control flow or white-box obfuscation.

6.5. Energy-Centric Systems

Participants: Axel Legay, Uli Fahrenberg.

This part is concerned with Tasks 1 and 2. Mostly, we focus on quantifying properties of interconnected objects such as Cyber Physical Systems (CPS) (SoS and CPS share a lot of commonalities).

Energy and resource management problems are important in areas such as embedded systems or autonomous systems. They are concerned with the question whether a given system admits infinite schedules during which (1) certain tasks can be repeatedly accomplished and (2) the system never runs out of energy (or other specified resources). Formal modeling and analysis of such problems has attracted some attention in recent years.

6.5.1. Papers:

- [18] (C; accepted) We define and study basic properties of $*$ -continuous Kleene ω -algebras that involve a $*$ -continuous Kleene algebra with a $*$ -continuous action on a semimodule and an infinite product operation that is also $*$ -continuous. We show that $*$ -continuous Kleene ω -algebras give rise to iteration semiring-semimodule pairs, and that for Büchi automata over $*$ -continuous Kleene ω -algebras, one can compute the associated infinitary power series.
- [17] (C; accepted) Energy problems are important in the formal analysis of embedded or autonomous systems. Using recent results on $*$ -continuous Kleene ω -algebras, we show here that energy problems can be solved by algebraic manipulations on the transition matrix of energy automata. To this end, we prove general results about certain classes of finitely additive functions on complete lattices which should be of a more general interest.
- [15] (C; accepted) We develop a $*$ -continuous Kleene ω -algebra of real-time energy functions. Together with corresponding automata, these can be used to model systems which can consume and regain energy (or other types of resources) depending on available time. Using recent results on $*$ -continuous Kleene ω -algebras and computability of certain manipulations on real-time energy functions, it follows that reachability and Büchi acceptance in real-time energy automata can be decided in a static way which only involves manipulations of real-time energy functions.

6.6. Languages for composition

Participants: Axel Legay, Thomas Given-Wilson.

This part is concerned with Task 1, especially to describe the composition of complex systems, and to study expressivity of existing formalisms.

Contemporary cyber-physical systems are inherently constructed out of a variety of agents with communication and interaction forming a key role in the behaviour of the system as a whole. Traditional approaches to reasoning over a single computation or treating the system as a single agent prove unsatisfactory for understanding the capabilities, strengths, and weaknesses of such systems.

Since communication is a fundamental to such systems it is necessary to understand the role the communication primitives themselves play. There are many approaches to communication primitives, often chosen for their ability to easily represent desired behaviour. However, the formal properties of many implementations or chosen models have not been presented.

An alternative to formalising each possible model individually is to abstract away and reason over families of models based on their communication primitives. This allows key results to be achieved in one model, and then generalised to the entire family, or transferred to other families based upon formal relations between these families. Thus making it possible for results to be easily applied to many models or systems without repeating significant effort.

6.6.1. Papers:

- [20] (C), [32] (J; submitted) The expressiveness of communication primitives has been explored in a common framework based on the π -calculus by considering four features: synchronism (asynchronous vs synchronous), arity (monadic vs polyadic data), communication medium (shared dataspaces vs channel-based), and pattern-matching (binding to a name vs testing name equality vs intensionality). Here another dimension coordination is considered that accounts for the number of processes required for an interaction to occur. Coordination generalises binary languages such as π -calculus to joining languages that combine inputs such as the Join Calculus and general rendezvous calculus. By means of possibility/impossibility of encodings, this paper shows coordination is unrelated to the other features. That is, joining languages are more expressive than binary languages, and no combination of the other features can encode a joining language into a binary language. Further, joining is not able to encode any of the other features unless they could be encoded otherwise.
- [33] (C; submitted) The expressiveness of communication primitives has been explored in a common framework by considering four features: synchronism, arity, communication medium, and pattern-matching. These all assume asymmetric communication between input and output primitives, however some calculi consider more symmetric approaches to communication such as fusion calculus and Concurrent Pattern Calculus. Symmetry can be considered either as allowing a mixture of input and output in an action or co-action, or as the unification of actions. By means of possibility/impossibility of encodings, this paper shows that: the action and co-action approach is related to or more expressive than many previously considered languages; and the unification approach is more expressive than some, but mostly unrelated to other languages.

GALAAD2 Team

6. New Results

6.1. Certifying isolated singular points and their multiplicity structure

Participant: Bernard Mourrain.

The paper [4] presents two new constructions related to singular solutions of polynomial systems. The first is a new deflation method for an isolated singular root. This construction uses a single linear differential form defined from the Jacobian matrix of the input, and defines the deflated system by applying this differential form to the original system. The advantages of this new deflation is that it does not introduce new variables and the increase in the number of equations is linear instead of the quadratic increase of previous methods. The second construction gives the coefficients of the so-called inverse system or dual basis, which defines the multiplicity structure at the singular root. We present a system of equations in the original variables plus a relatively small number of new variables. We show that the roots of this new system include the original singular root but now with multiplicity one, and the new variables uniquely determine the multiplicity structure. Both constructions are "exact", meaning that they permit one to treat all conjugate roots simultaneously and can be used in certification procedures for singular roots and their multiplicity structure with respect to an exact rational polynomial system.

Joint work with Agnes Szanto, Department of Mathematics, North Carolina State University, Raleigh, USA; Jonathan D. Hauenstein, Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, USA.

6.2. On the construction of general cubature formula by flat extensions

Participants: Marta Abril-Bucero, Bernard Mourrain.

We describe a new method to compute general cubature formulae [5]. The problem is initially transformed into the computation of truncated Hankel operators with flat extensions. We then analyse the algebraic properties associated to flat extensions and show how to recover the cubature points and weights from the truncated Hankel operator. We next present an algorithm to test the flat extension property and to additionally compute the decomposition. To generate cubature formulae with a minimal number of points, we propose a new relaxation hierarchy of convex optimization problems minimizing the nuclear norm of the Hankel operators. For a suitably high order of convex relaxation, the minimizer of the optimization problem corresponds to a cubature formula. Furthermore cubature formulae with a minimal number of points are associated to faces of the convex sets. We illustrate our method on some examples, and for each we obtain a new minimal cubature formula.

This is a joint work with C. Bajaj (Univ. of Austin, Texas, USA).

6.3. A moment matrix approach to computing symmetric cubatures

Participants: Mathieu Collowald, Evelyne Hubert.

A quadrature is an approximation of the definite integral of a function by a weighted sum of function values at specified points, or nodes, within the domain of integration. Gaussian quadratures are constructed to yield exact results for any polynomial of degree $2r - 1$ or less by a suitable choice of r nodes and weights. Cubature is a generalization of quadrature in higher dimension. Constructing a cubature amounts to find a linear form

$$\Lambda : \mathbb{R}[x] \rightarrow \mathbb{R}, p \mapsto \sum_{j=1}^r a_j p(\xi_j)$$

from the knowledge of its restriction to $\mathbb{R}[x]_{\leq d}$. The unknowns are the number of nodes r , the weights a_j and the nodes ξ_j .

In [7] we use a basis-free version of an approach to cubatures based on moment matrices in terms of the Hankel operator \mathcal{H} associated to Λ . The existence of a cubature of degree d with r nodes boils down to conditions of ranks and positive semidefiniteness on \mathcal{H} . We then recognize the nodes as the solutions of a generalized eigenvalue problem.

Standard domains of integration are symmetric under the action of a finite group. It is natural to look for cubatures that respect this symmetry. Introducing adapted bases obtained from representation theory, the symmetry constraint allows to block diagonalize the Hankel operator \mathcal{H} . We then deal with smaller-sized matrices both for securing the existence of the cubature and computing the nodes. The sizes of the blocks are furthermore explicitly related to the orbit types of the nodes with the new concept of the matrix of multiplicities of a finite group. It provides preliminary criteria of existence of a cubature with a given organisation of the nodes in orbit types.

The Maple implementation of the presented algorithms allows to determine, with moderate computational efforts, all the symmetric cubatures of a given degree. We present new relevant cubatures.

6.4. Invariantization of symmetric polynomial systems

Participants: Mathieu Collowald, Evelyne Hubert.

Assuming the variety of a set of polynomials is invariant under a group action, we provide a set of invariants that define the same variety. The contribution is about infinite algebraic groups, the case of finite group being previously known. We introduce for those a new concept of algebraic invariantization. It is based on the construction of rational invariants by Hubert and Kogan [14], a construction for which we provide here new simplified proofs.

6.5. Effective criterions for bigraded birational maps

Participant: Laurent Busé.

A rational map $\mathcal{F} : \mathbb{P}^m \dashrightarrow \mathbb{P}^n$ between projective spaces is defined by a collection of homogeneous polynomials $\mathbf{f} := (f_0, \dots, f_n)$ in $m + 1$ variables of the same degree. The problem of deciding or providing sufficient conditions for such a map \mathcal{F} to be birational have attracted a lot of interest in the past and it is still an active area of research. Methods that are based of some properties of the syzygies of \mathbf{f} are definitely the more adapted for computational purposes in the sense that they make the problem of birationality effectively computable in the usual implementation of the Gröbner basis algorithm. The goal of this work is to extend these syzygies-based methods and techniques to the context of rational maps whose source is a product of two projective spaces $\mathbb{P}^r \times \mathbb{P}^s$ instead.

An important motivation for considering bi-graded rational maps comes from the field of geometric modeling. Indeed, the geometric modeling community uses almost exclusively bi-graded rational maps for parameterizing curves, surfaces or volumes under the name of rational tensor-product Bézier parameterizations. It turns out that an important property is to guarantee the birationality of these parameterizations onto their images. An even more important property is to preserve this birationality property during a design process, that is to say when the coefficients of the defining polynomials are continuously modified. As a first attempt to tackle these difficult problems, we analyze in detail birational maps from $\mathbb{P}^1 \times \mathbb{P}^1$ to \mathbb{P}^2 in low bi-degree by means of syzygies.

This work is done in the context of the SYRAM project which is funded by the MathAmSud programme. It is a collaboration with N. Botbol (University of Buenos Aires), M. Chardin (University of Paris 6), H. Hassanzadeh (University of Rio de Janeiro), A. Simis (University de Pernambuco) and Q. H. Tran (University of Paris 6). A paper is in preparation.

6.6. Orthogonal projection of points on Bézier curves and surfaces

Participant: Laurent Busé.

In this work, we introduce a new method for computing the orthogonal projections of a point onto a Bézier curve or surface. It is based on the concept of matrix representation we have introduced and developed in some previous works, which is here applied to the parameterizations of the normal planes or lines of a curve or surface, respectively. It consists in the computation of a matrix depending of the ambient space variables, which is done in a pre-processing step, and then the use of tools from numerical linear algebra for a fast and accurate solving of each instance of the problem.

This is an on going work done in the context of the SYRAM project which is funded by the MathAmSud programme. It is a collaboration with N. Botbol (University of Buenos Aires) and M. Chardin (University of Paris 6).

6.7. Extraction of cylinders and cones from minimal point sets

Participants: Laurent Busé, André Galligo, Jiajun Zhang.

The extraction of geometric primitives from 3D point clouds is an important problem in reverse engineering. These 3D point clouds are typically obtained by means of accurate 3D scanners and there exists several methods for performing the 3D geometric primitives extraction. An important category among these methods are based on a RANSAC method. For such methods, the primitives are directly extracted for the input point cloud. The basic idea is to extract a particular elementary type of shape, such as planes, spheres, cylinders, cones or tori, from the smallest possible set of points and then to judge if this extracted primitive is relevant to the full point cloud. Therefore, for this category of methods it is very important to compute a particular type of shape through the smallest possible number of points, including normals or not. If the extraction of planes and spheres is easy to treat, the cases of cylinders, cones and tori are more involved. In this work, we aim at developing methods for extracting these geometric primitives from the smaller possible number of points (counting multiplicities if normals are taken into account). Another objective is also to provide methods for extraction without using estimated normals in order to improve the accuracy of the extracted geometric primitive, or to use mixed data depending of the applied context (some points with normals and some other points without normals). A paper is in preparation.

6.8. Discriminant of a complete intersection space curve

Participant: Laurent Busé.

In this work, we develop the formalism of the discriminant of a complete intersection curve in the three dimensional projective space, that is to say a curve which is represented as the zero locus of two homogeneous polynomials in four variables. Our main objective is to provide a new computational approach to this object without relying on the so-called "Cayley trick" for which it is necessary to introduce new variables. We also aim at getting a universal definition of this discriminant over the integers so that it holds under any specialization of the coefficients to an arbitrary commutative ring. Another aspect of this work is to explore properties of this discriminant, typically invariance, covariance and change of basis properties.

This is an on going work which is done in collaboration with Ibrahim Nonkane (University of Ouagadougou, Burkina Faso).

6.9. Resultants, flexes, and the generalization of Salmon's formula

Participant: Laurent Busé.

Given an algebraic variety $S \subset \mathbb{P}^n$ and a point $p \in S$, the osculation order of the point p is the maximum of the multiplicity of intersection at p of S with any line through p . We denote it by μ_p and define $Flex(S) = \{p \in \mathbb{P}^n | \mu_p > n\}$.

If $n = 2$, it is known that if C is a plane algebraic curve of degree d then $Flex(C)$ is the intersection of C with its Hessian, this latter being of degree $3d - 6$. A famous generalization of this result to the case $n = 3$ has been obtained by Salmon in 1860: for a general variety S , $Flex(S)$ is the intersection of S with another hypersurface of degree $11d - 24$. In this work, we are studying the generalization of this formula to arbitrary dimension n . We proved that given $S \subset \mathbb{P}^n$ of degree d , $Flex(S)$ is obtained by intersecting S with another hypersurface of degree

$$d \left(\sum_{k=1}^n \frac{n!}{k} \right) - n!$$

We are also looking for an explicit expression of an equation of this latter hypersurface.

This is a work in progress which is done in the context of a PICS collaboration funded by CNRS. It is a joint work with M. Chardin (University Paris 6), C. D'Andrea (University of Barcelona), M. Sombra (University of Barcelona) and M. Weiman (University of Caen).

6.10. Computer Algebra Applied to a Solitary Waves Study

Participant: André Galligo.

In [3], we apply Computer algebra techniques, such as algebraic computations of resultants and discriminants, certified drawing (with a guaranteed topology) of plane curves, to a problem in Fluid dynamics: We investigate “capillary-gravity” solitary waves in shallow water, relying on the framework of the Serre-Green-Naghdi equations. So, we deal with 2 dimensional surface waves, propagating in a shallow water of constant depth. By a differential elimination process, the study reduces to describing the solutions of an ordinary non linear first order differential equation, depending on two parameters. The paper is illustrated with examples and pictures computed with the computer algebra system Maple.

Joint work with Didier Clamond (University of Nice, France) and Denys Dutykh (University of Le Bourget, France).

6.11. H1-parameterizations of plane physical domains with complex topology in Isogeometric analysis

Participants: André Galligo, Bernard Mourrain, Meng Wu.

Isogeometric analysis (IGA) is a method for solving geometric partial differential equations(PDEs). Generating parameterizations of a PDE's physical domain is a basic and important issue within IGA framework. In [13], we present a global H1-parameterization method for a planar physical domain with complex topology.

Joint work with B. NKonga, Univeristy of Nice - Sophia Antipolis and EPI CASTOR, Inria.

GALLIUM Project-Team

7. New Results

7.1. Formal verification of compilers and static analyzers

7.1.1. *The CompCert formally-verified compiler*

Participants: Xavier Leroy, Jacques-Henri Jourdan, François Pottier, Bernhard Schommer [AbsInt GmbH].

In the context of our work on compiler verification (see section 3.3.1), since 2005 we have been developing and formally verifying a moderately-optimizing compiler for a large subset of the C programming language, generating assembly code for the PowerPC, ARM, and x86 architectures [6]. This compiler comprises a back-end, which translates the Cminor intermediate language to PowerPC assembly and is reusable for source languages other than C [5], and a front-end, which translates the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq's extraction facility generates executable OCaml code. The compiler comes with a 50000-line, machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, we improved the CompCert C compiler in several directions:

- The generation of debugging information in DWARF format was implemented by Bernhard Schommer at AbsInt. Consequently, CompCert-compiled programs can now be debugged using standard debuggers. Xavier Leroy extended the back-end compilation passes and their proofs to propagate debugging information throughout the compilation pipeline.
- The CompCert formal semantics was made more precise in order to increase confidence. We tightened the semantics of pointer comparisons against the null pointer. We formalized the distinction between public and private (`static`) global definitions, and used it to prove the correctness of the “Unusedglob” pass that removes unreferenced private definitions.
- The calling conventions used to pass function arguments and results of `struct` and `union` types were revised in order to comply with the Application Binary Interfaces of the target platforms.
- We added partial support for extended inline assembly, an extension of the C language popularized by the GCC compiler and often used in low-level code.
- Detailed explanations of syntax errors are now produced. This usability feature builds on François Pottier's work on error reporting in LR parsers (see section 7.4.4).
- The PowerPC back-end was extended to support the PowerPC 64-bit extensions and the Freescale E5500 variant.

We released two versions of CompCert, integrating these enhancements: version 2.5 in June and version 2.6 in December. This is the public version of CompCert, available for evaluation and research purposes. In parallel, our industrial partner, **AbsInt Angewandte Informatik GmbH**, sells a commercial version of CompCert with long-term maintenance.

7.1.2. *Formal verification of static analyzers based on abstract interpretation*

Participants: Jacques-Henri Jourdan, Xavier Leroy, Sandrine Blazy [team Celtique], Vincent Laporte [team Celtique], David Pichardie [team Celtique], Sylvain Boulmé [Grenoble INP, VERIMAG], Alexis Foulhé [Université Joseph Fourier de Grenoble, VERIMAG], Michaël Périn [Université Joseph Fourier de Grenoble, VERIMAG].

In the context of the ANR Verasco project, we are investigating the formal specification and verification in Coq of a realistic static analyzer based on abstract interpretation. This static analyzer handles a large subset of the C language (the same subset as the CompCert compiler, minus recursion and dynamic allocation); supports a combination of abstract domains, including relational domains; and should produce usable alarms. The long-term goal is to obtain a static analyzer that can be used to prove safety properties of real-world embedded C code. The overall architecture and specification of Verasco is described in a paper that was presented at POPL 2015 [19].

This year, Jacques-Henri Jourdan continued the development of this static analyzer, with two goals. First, Jacques-Henri Jourdan improved the precision and analysis time of the existing abstract domains. The existing communication system between domains was instantiated to the cooperation between the abstract domain of intervals and the abstract domain of congruences. Second, Jacques-Henri Jourdan implemented and formalized in our static analyzer the Octagon abstract domain of Miné [46]. This led to new results in the theory behind this abstract domain, allowing Jourdan to use sparse data structures for representing octagons.

7.1.3. A SPARK Front-end for CompCert

Participants: Pierre Courtieu, Zhi Zang [Kansas University].

SPARK is a language, and a platform, dedicated to developing and verifying critical software. It is a subset of the Ada language. It shares with Ada a strict typing discipline and gives strict guarantees in terms of safety. SPARK goes one step further by disallowing certain “dangerous” features, that is, those that are too difficult to statically analyze (aliasing, references, etc). Given its dedication to safety critical software, we think that the SPARK platform can benefit from a certified compiler. We are working on adding a SPARK front-end to the CompCert verified compiler.

Defining a semantics for SPARK in Coq is previous joint work with Zhi Zang from Kansas University. The current front-end is based on this semantics. The compiler has been written and tested, and the proofs of correctness are currently under way.

7.1.4. Verified JIT compilation of Coq

Participants: Maxime Dénès, Xavier Leroy.

Last year, we started the Coqonut project, whose objective is to develop and formally verify an efficient, compiled implementation of Coq’s reduction. This year, we made progress on this verification effort:

- We ported our OCaml prototype to Coq and started its verification, notably of the first phase of the compiler which involves uncurrying, using untyped step-indexed logical relations.
- We adapted (part of) the Coq x86 macro assembler by Andrew Kennedy, Nick Benton, Jonas B. Jensen and Pierre-Evariste Dagand to x86-64. This macro assembler framework is used in Coqonut’s backend to generate assembly or machine code.

7.2. Language design and type systems

7.2.1. Full reduction in the presence of inconsistent assumptions

Participants: Didier Rémy, Gabriel Scherer.

Gabriel Scherer and Didier Rémy continued their work on assumption hiding and presented it at ESOP 2015 [22]. This work aims at restoring confluence when mixing full and weak reduction and providing a continuum between consistent and inconsistent abstraction. Assumption hiding supports fine-grained control of dependencies between computations and the logical hypotheses they depend on. Although studied for a language of coercions, the solution is more general and should be applicable to any language with abstraction over propositions that are left implicit, either for the user’s convenience in a surface language or because they have been erased prior to computation in an internal language.

7.2.2. Equivalence and normalization of lambda-terms with sums

Participants: Gabriel Scherer, Guillaume Munch-Maccagnoni [Université Paris-Diderot, laboratoire PPS].

Gabriel Scherer presented at TLCA 2015 his work on understanding equivalence of sum types using the proof-theoretical technique of focusing [24]. Independently, his collaboration with Guillaume Munch-Maccagnoni resulted in a presentation of sum equivalence using an abstract machine calculus [33]. This approach allows for a more concise and cleaner definition of the equivalence relation, and a finer-grained understanding of the role of purity assumptions in the program equivalence relation.

7.2.3. *Types with unique inhabitants for code inference*

Participants: Gabriel Scherer, Didier Rémy.

Gabriel Scherer and Didier Rémy presented at ICFP 2015 [23] an algorithm to decide whether a type has a unique inhabitant in the simply-typed lambda-calculus with sum types. This algorithm comes along with a prototype implementation. This minimal setting is not representative of the expressiveness of realistic programming languages, but already covers a first few interesting code inference scenarios for polymorphic libraries in functional languages with prenex polymorphism: for instance, we can infer the “bind” function of the exception monad.

7.2.4. *Refactoring with ornaments in ML*

Participants: Thomas Williams, Didier Rémy.

Thomas Williams and Didier Rémy continued working on ornaments for program refactoring and program transformation in ML. Ornaments have been introduced as a way to describe some changes in data type definitions that preserve their recursive structure, reorganizing, adding, or dropping some pieces of data. After a new data structure has been described as an ornament of an older one, some functions operating on the bare structure can be partially or sometimes totally lifted into functions operating on the ornamented structure.

We have previously described an algorithm to perform this lifting in ML. This description was informal. This year, we improved this algorithm by decomposing it in several steps and we formalized it. Using ornament inference, we first elaborate an ML program into a generic program, which can be seen as a template for all possible liftings of the original program. The generic program is defined in a superset of ML. It can then be instantiated with specific ornaments, and simplified back into an ML program. We also studied the properties of lifting, particularly the preservation of complexity and effects, with the aim of characterizing more precisely the syntactic liftings that can be produced by our algorithm.

On the practical side, our prototype ornamentation tool has been improved with an implementation of ornament inference. The generalized program gives a description of all possible extension points that must be filled by providing patches. In practice, a few heuristics are enough to automate most of the patching work. The rest can be filled interactively by the programmer. In the case of refactoring (the representation of a data type is modified without adding any data), the transformation is fully automatic.

7.2.5. *The Mezzo programming language*

Participants: Thibaut Balabonski [Université Paris Sud], François Pottier, Jonathan Protzenko.

Mezzo is a programming language proposal whose untyped foundation is very much like OCaml (i.e., it is equipped with higher-order functions, algebraic data structures, mutable state, and shared-memory concurrency) and whose type system offers flexible means of describing ownership policies and controlling side effects.

A comprehensive paper, which contains both a tutorial introduction to Mezzo and a description of its formal definition and proof, was submitted to TOPLAS in 2014. This year, after a round of reviewing, it was revised and accepted for publication [11]. A reflection on the design of Mezzo was presented at SNAPL 2015 [21].

7.3. Shared-memory parallelism

7.3.1. *Weak memory models*

Participants: Luc Maranget, Jade Alglave [Microsoft Research, Cambridge], Patrick Cousot [New York University], Keryan Didier.

Modern multi-core and multi-processor computers do not follow the intuitive “Sequential Consistency” model that would define a concurrent execution as the interleaving of the executions of its constituent threads and that would command instantaneous writes to the shared memory. This situation is due both to in-core optimisations such as speculative and out-of-order execution of instructions, and to the presence of sophisticated (and cooperating) caching devices between processors and memory. Luc Maranget took part in an international research effort to define the semantics of the computers of the multi-core era, and more generally of shared-memory parallel devices or languages, with a clear focus on devices.

More precisely, in 2015, Luc Maranget collaborated with Jade Alglave and Patrick Cousot to extend “Cats”, a domain-specific language for defining and executing weak memory models. A precise semantics for “Cats” is the core of a submitted journal article that also includes a study and formalisation of the HSA memory model — the Heterogeneous System Architecture foundation is an industry standards body targeting heterogeneous computing devices (see <http://www.hsafoundation.com/>). The new extensions of the Cats language have been integrated in the released version of the **diy** tool suite (see section 6.2).

Luc Maranget also co-authored a paper that will be presented at POPL 2016 [18]. This work describes an operational semantics for the new generation ARM processors. It is joint work with many researchers, including S. Flur and other members of P. Sewell’s team (University of Cambridge) and W. Deacon (ARM Ltd).

During his M2 internship, supervised by Luc Maranget, Keryan Didier significantly improved the **diy** tool suite, in particular by writing front-ends for ARMv8 and for a subset of the C language. Keryan Didier also wrote a new (as yet unreleased) tool to translate between various input languages, in particular from machine assemblers to generic assembler and back.

7.3.2. Algorithms and data structures for parallel computing

Participants: Umut Acar, Vitalii Aksenov, Arthur Charguéraud, Mike Rainey, Filip Sieczkowski.

The ERC Deepsea project, with principal investigator Umut Acar, started in June 2013 and is hosted by the Gallium team. This project aims at developing techniques for parallel and self-adjusting computation in the context of shared-memory multiprocessors (i.e., multicore platforms). The project is continuing work that began at Max Planck Institute for Software Systems between 2010 and 2013. As part of this project, we are developing a C++ library, called PASL, for programming parallel computations at a high level of abstraction. We use this library to evaluate new algorithms and data structures. We obtained three major results this year.

Our result on the development of fast and robust parallel graph traversal algorithms based on depth-first-search has been presented at the ACM/IEEE Conference on High Performance Computing [15]. This algorithm leverages a new sequence data structure for representing the set of edges remaining to be visited. In particular, it uses a balanced split operation for partitioning the edges of a graph among the processors involved in the computation. Compared with prior work, the new algorithm is designed to be efficient not just for particular classes of graphs, but for all input graphs.

Our second result is a calculus for parallel computing on hardware shared memory computers such as modern multicores. Many languages for writing parallel programs have been developed. These languages offer several distinct abstractions for parallelism, such as fork-join, async-finish, futures, etc. While they may seem similar, these abstractions lead to different semantics, language design and implementation decisions. In this project, we consider the question of whether it would be possible to unify these approaches to parallelism. To this end, we propose a calculus, called the *DAG-calculus*, which can encode existing approaches to parallelism based on fork-join, async-finish, and futures, and possibly others. We have shown that the approach is realistic by presenting an implementation in C++ and by performing an empirical evaluation. This work has been submitted for publication.

Our third result concerns the development of parallel dynamic algorithms. This year, we started developing a parallel dynamic algorithm for tree computations. The algorithm is dynamic in the sense that it admits changes to the underlying tree in the form of insertions and deletions of edges and vertices and updates the computation by doing total work that is linear in the size of the changes, but only logarithmic in the size of the tree. The

algorithm is parallel in the sense that the updates take place in parallel. Parallel algorithms have been studied extensively in the past, but few of these are dynamic. Similarly, dynamic algorithms have also been studied extensively in the past, but few of these are parallel. Our work thus explores what in retrospect seems like an obvious gap in the literature. A paper describing this work is in preparation.

7.4. The OCaml language and system

7.4.1. The OCaml system

Participants: Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant, Xavier Leroy, Luc Maranget, Gabriel Scherer, Mark Shinwell [Jane Street], Leo White [Jane Street], Jeremy Yallop [OCaml Labs, Cambridge University].

This year, we released versions 4.02.2 and 4.02.3 of the OCaml system. These are minor releases that fix about 100 bugs and implement 12 minor new features, including support for nonrecursive type definitions and a higher-level interface with documentation generation tools.

Most of our activity was devoted to preparing the next major release of OCaml, version 4.03.0, which is expected in the first quarter of 2016. The novelties we worked on include:

- Inline record types as arguments to constructors of sum types, combining the clarity and extensibility brought by named record fields with the compact in-memory representation of unnamed constructor arguments.
- Improved redundancy and exhaustiveness checks for pattern-matching over generalized algebraic data types (GADTs) [41].
- Improved unboxing optimizations for numbers, including the ability to mark arguments and results of external C functions as unboxed.
- The garbage collector was made more incremental, so as to reduce the worst-case GC pause times.
- The native-code compiler was ported to two new architectures: PowerPC 64 bits (including IBM's new little-endian variant) and IBM zSystems.

On the organization side, we switched to Github as the central repository for the OCaml development sources. Github facilitates collaborative work among the growing community of contributors to the OCaml code base. In 2015, more than 100 contributors proposed small or large improvements to the OCaml compiler distribution.

7.4.2. Memory profiling OCaml applications

Participants: Fabrice Le Fessant, Çagdas Bozman [OCamlPro], Albin Coquereau [OCamlPro].

Most modern languages make use of automatic memory management to discharge the programmer from the burden of explicitly allocating and releasing chunks of memory. As a consequence, when an application exhibits an unexpected usage of memory, programmers need new tools to understand what is happening and how to solve such an issue. In OCaml, the compact representation of values, with almost no runtime type information, makes the design of such tools more complex.

In the past, we have experimented with different tools to profile the memory usage of real OCaml applications, in particular one that saves snapshots of the heap after every garbage collection. Snapshots can then be analysed to display the evolution of memory usage, with detailed information on the types of values, where they were allocated and from where they are still reachable.

This year, we experimented in three new directions, mostly driven by the size of the snapshots to be analysed:

- We studied several ways of displaying snapshots. Because of the large amount of information contained in a snapshot, it is hard for a typical user to find what he or she is looking for. We tried multiple filtering methods, based on graph algorithms, to remove the least significant information from the reports given to the user.

- We experimented with new algorithms to compress and analyse *huge* memory snapshots, i.e., snapshots that are too big to fit in the computer’s memory. Indeed, standard analyses on snapshots bigger than the available memory are too long to run in practice because of random disk accesses. Thus, we tried several compression methods for snapshots and graph-reduced them to fit in memory, without losing any information, reaching a 50x speedup in complete analysis time.
- We implemented a new graph algorithm to merge sets of blocks in memory by the sets of roots they are reachable from. Such a computation was heretofore supposed to be untractable in practice, but could actually be computed in our case on huge compressed snapshots in reasonable time.

7.4.3. Advanced development tools for OCaml

Participants: Fabrice Le Fessant, Pierre Chambart [OCamlPro], Michael Laporte [OCamlPro].

In order to promote the use of OCaml in industrial contexts, we have worked on improving the tools that accompany OCaml:

- We developed the first prototype of a native debugger for OCaml, based on the LLDB debugging framework on top of LLVM. For that, we first generated a full OCaml binding for the LLDB library, by parsing the C++ headers of the libraries and automatically generating OCaml and C++ stubs. We were then able to use the OCaml binding to develop several tools, ranging from a simple tool that displays the internal GC information of a finished OCaml application, to an almost complete debugger, which displays OCaml values using runtime type information added for memory profiling.
- We also developed a new profiling framework for OCaml, called *operf*. The framework is composed of two tools: *operf-micro* can be used to run micro-benchmarks directly from inside modified OCaml compiler sources, while the *operf-macro* tool can be used to evaluate the impact of a new compiler optimization on a large set of OPAM packages.
- Finally, we came up with new ideas for *ocp-build*, a generic building tool with OCaml-specific support, to improve the expressiveness of its package description language and to easily describe cross-compilation of OCaml packages.

7.4.4. Error diagnosis in Menhir parsers

Participant: François Pottier.

LR parsers are powerful and efficient, but traditionally have done a poor job of explaining syntax errors. Although it is easy to report where an error was detected, it seems difficult to explain what has been understood so far and what is expected next. The OCaml and CompCert compilers, until now, have offered little information to the user beyond the traditional “syntax error” message.

In 2003, Jeffery proposed associating a fixed diagnostic message with every state of the LR automaton (therefore ignoring the automaton’s stack). This simple approach may seem tempting. However, a typical automaton has hundreds or thousands of states. Not all of them can trigger an error, but it is difficult to tell which can, and which cannot. Furthermore, for certain states, it is difficult (or even impossible) to write an accurate diagnostic message, because some vital contextual information resides in the stack, which Jeffery’s method cannot access.

In 2015, François Pottier proposed a reachability algorithm for LR automata, which he implemented in the Menhir parser generator (see section 6.3). This algorithm allows finding out which states can trigger an error and (therefore) require writing a diagnostic message. Furthermore, Pottier proposed two mechanisms for influencing where errors are detected. If used appropriately, these mechanisms make it easier (or possible) to write an accurate diagnostic message.

Pottier applied this approach to the C grammar in the front-end of the CompCert compiler, therefore allowing CompCert to produce better diagnostic messages when a C program is syntactically incorrect.

A short paper describing this work will be presented at JFLA 2016 [29]. A longer paper is in submission.

7.4.5. Improvements to Menhir

Participants: Frédéric Bour [independent consultant], Jacques-Henri Jourdan, François Pottier, Yann Régis-Gianas [team πr^2], Gabriel Scherer.

In 2015, The Menhir parser generator (see section 6.3) was extended with many new features, several of which originated in the Merlin IDE for OCaml and were ported back into Menhir.

- The parsers generated by Menhir are now incremental: they can be stopped and resumed at any point, at essentially no cost. This is exploited in Merlin, where the text is re-parsed after every keystroke.
- The state of the parser can be inspected by the user. This allows building custom libraries, outside Menhir, for error diagnosis, error recovery, etc. This is exploited in Merlin, where a valid abstract syntax tree is built (and passed to the OCaml type-checker) even if the text contains syntax errors.
- A reachability algorithm has been implemented (see section 7.4.4). It allows finding out which states can trigger an error and (therefore) require a diagnostic message to be written. It is accompanied with several tools that help maintain the database of diagnostic messages as the grammar evolves.
- Compatibility with `ocamlyacc` has been improved, in particular insofar as the computation of locations is concerned. This should help port the OCaml parser from `ocamlyacc` to Menhir, a transition that we envision making in the near future. This should help improve the quality of OCaml's syntax error messages.

7.5. Software specification and verification

7.5.1. Machine-checked proofs of programs, including time complexity

Participants: Arthur Charguéraud, Armaël Guéneau, François Pottier.

In a security-critical setting, it is important to prove that a program is correct, and to do so formally, that is, via a machine-checked proof. It is also important, one may argue, to prove that the program does not require more resources than expected (where a “resource” may be time, memory space, disk space, network bandwidth, etc.). Otherwise, even though the program is “correct” in theory, it may turn out to be unusable in practice.

Separation Logic, extended with the notion of a “time credit”, a permission to perform one step of computation, allows reasoning about the correctness and the (amortized) time complexity of a program. Using this approach, which Charguéraud implemented in the CFML tool, Charguéraud and Pottier produced a machine-checked proof of the correctness and time complexity of a Union-Find data structure, implemented as an OCaml module. This demonstrates that this approach scales up to difficult complexity analyses and down to the level of actual executable code (as opposed to pseudo-code). This work was presented at ITP 2015 [17].

During his M2 internship, Armaël Guéneau extended this approach so as to allow working conveniently with the big- O notation. He extended the CFML library and verified the time complexity of a binary random access list data structure due to Okasaki. This work has not been published yet.

7.5.2. Verified property-based random testing

Participants: Zoe Paraskevopoulou [ENS Cachan, team Prosecco], Cătălin Hrițcu [team Prosecco], Maxime Dénès, Leonidas Lampropoulos [U. of Pennsylvania], Benjamin C. Pierce [U. of Pennsylvania].

Property-based random testing has been popularized in the functional programming community by tools like QuickCheck. Its integration with a proof assistant creates an interesting opportunity: reusable or tricky testing code can be formally verified using the proof assistant itself.

We introduced a novel methodology for formally verified property-based testing and implemented it as a foundational verification framework for QuickChick, a port of QuickCheck to Coq. Our framework enables one to verify that the executable testing code is testing the right Coq property. To make verification tractable, we provided a systematic way for reasoning about the set of outcomes a random data generator can produce with non-zero probability, while abstracting away from the actual probabilities.

We also applied this methodology to a complex case study on testing an information-flow control abstract machine, demonstrating that our verification methodology is modular and scalable and that it requires minimal changes to existing code.

Maxime Dénès more specifically contributed to the development of the QuickChick Coq plug-in, to the development of Coq libraries for reasoning on the set of outcomes of random generators and to the verification of QuickChick's combinator library.

This work was presented at ITP 2015 [20].

7.5.3. *Tools for TLA+*

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Martin Riener [team VeriDis], Stephan Merz [team VeriDis].

Damien Doligez is head of the “Tools for Proofs” team in the Microsoft-Inria Joint Centre. The aim of this project is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing Lamport's ideas [43], and to build tools for writing TLA+ specifications and mechanically checking the proofs.

This year, we released version 1.4.3 of the TLA+ Proof System (TLAPS) [40], the part of the TLA+ tools that handles mechanical checking of TLA+ proofs.

This was the last year of the ADN4SE project, which develops tools for rapid development of real-time software based on the PharOS real-time kernel developed by CEA. Within this project we built, in collaboration with CEA, a formal proof of determinacy of the message-passing subsystem of PharOS. We used this experience to improve our TLA+ tools and libraries.

We have started a rewrite of TLAPS from scratch, which will make it possible to handle all aspects of the TLA+ language, including temporal formulas and their proofs.

7.5.4. *Certified distributed algorithms for autonomous mobile robots*

Participants: Pierre Courtieu, Xavier Urbain [ENSIIE], Sébastien Tixeuil [U. Pierre et Marie Curie], Lionel Rieg [Collège de France].

The variety and complexity of the tasks that can be performed by autonomous robots are increasing. Many applications envision groups of mobile robots that self-organise and cooperate toward the resolution of common objectives, in the absence of any central coordinating authority.

We are developing a Coq-based verification platform for distributed algorithms for autonomous robots. This year, we mechanically proved and slightly generalized a non-trivial proof of impossibility of such an algorithm under certain hypotheses [14]. We also proved several algorithms in the literature, demonstrating the viability of the platform [13].

7.5.5. *Contributions to ProofGeneral, an IDE for Coq*

Participant: Pierre Courtieu.

User interface is a crucial issue for theorem provers like Coq. ProofGeneral [38], an emacs-based prover interface, is widely used among Coq users. In addition to synchronizing with the evolutions of Coq itself, we contributed many improvements to ProofGeneral during the past year, among which: a better debugging mode and message printing, user assistance for naming hypotheses and indenting proof scripts, and more.

GEOMETRICA Project-Team

7. New Results

7.1. Mesh Generation and Geometry processing

7.1.1. *Discrete Derivatives of Vector Fields on Surfaces An Operator Approach*

Participants: Frédéric Chazal, Maksim Ovsjanikov.

In collaboration with O. Azencot, M. Ben Chen (Technion, Israel Institute of Technology).

Vector fields on surfaces are fundamental in various applications in computer graphics and geometry processing. In many cases, in addition to representing vector fields, the need arises to compute their derivatives, for example, for solving partial differential equations on surfaces or for designing vector fields with prescribed smoothness properties. In this work, we consider the problem of computing the Levi-Civita covariant derivative, that is, the tangential component of the standard directional derivative, on triangle meshes. This problem is challenging since, formally, tangent vector fields on polygonal meshes are often viewed as being discontinuous, hence it is not obvious what a good derivative formulation would be. We leverage the relationship between the Levi-Civita covariant derivative of a vector field and the directional derivative of its component functions to provide a simple, easy-to-implement discretization for which we demonstrate experimental convergence. In addition, we introduce two linear operators which provide access to additional constructs in Riemannian geometry that are not easy to discretize otherwise, including the parallel transport operator which can be seen simply as a certain matrix exponential. Finally, we show the applicability of our operator to various tasks, such as fluid simulation on curved surfaces and vector field design, by posing algebraic constraints on the covariant derivative operator.

7.1.2. *Isotopic Meshing within a Tolerance Volume*

Participant: David Cohen-Steiner.

In collaboration with M. Mandad, P. Alliez (Titane Project-team).

We give an algorithm [22] that generates from an input tolerance volume a surface triangle mesh guaranteed to be within the tolerance, intersection free and topologically correct. A pliant meshing algorithm is used to capture the topology and discover the anisotropy in the input tolerance volume in order to generate a concise output. We first refine a 3D Delaunay triangulation over the tolerance volume while maintaining a piecewise-linear function on this triangulation, until an isosurface of this function matches the topology sought after. We then embed the isosurface into the 3D triangulation via mutual tessellation, and simplify it while preserving the topology. Our approach extends to surfaces with boundaries and to non-manifold surfaces. We demonstrate the versatility and efficacy of our approach on a variety of data sets and tolerance volumes.

7.1.3. *CGALmesh: A Generic Framework for Delaunay Mesh Generation*

Participants: Jean-Daniel Boissonnat, Clément Jamin, Mariette Yvinec.

In collaboration with P. Alliez (Titane Project-team).

CGALmesh [21] is the mesh generation software package of the Computational Geometry Algorithm Library (CGAL). It generates isotropic simplicial meshes—surface triangular meshes or volume tetrahedral meshes—from input surfaces, 3D domains, and 3D multidomains, with or without sharp features. The underlying meshing algorithm relies on restricted Delaunay triangulations to approximate domains and surfaces and on Delaunay refinement to ensure both approximation accuracy and mesh quality. CGALmesh provides guarantees on approximation quality and on the size and shape of the mesh elements. It provides four optional mesh optimization algorithms to further improve the mesh quality. A distinctive property of CGALmesh is its high flexibility with respect to the input domain representation. Such a flexibility is achieved through a careful software design, gathering into a single abstract concept, denoted by the oracle, all required interface features between the meshing engine and the input domain. We already provide oracles for domains defined by polyhedral and implicit surfaces.

7.2. Topological and Geometric Inference

7.2.1. *Subsampling Methods for Persistent Homology*

Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with B.T. Fasy, F. Lecci, A. Rinaldo and L. Wasserman (Carnegie Mellon University).

Persistent homology is a multiscale method for analyzing the shape of sets and functions from point cloud data arising from an unknown distribution supported on those sets. When the size of the sample is large, direct computation of the persistent homology is prohibitive due to the combinatorial nature of the existing algorithms. We propose to compute the persistent homology of several subsamples of the data and then combine the resulting estimates. We study the risk of two estimators and we prove that the subsampling approach carries stable topological information while achieving a great reduction in computational complexity.

7.2.2. *Efficient and Robust Persistent Homology for Measures*

Participants: Frédéric Chazal, Steve Oudot.

In collaboration with M. Buchet (Ohio State University) and Donald Sheehy (University of Connecticut).

A new paradigm for point cloud data analysis has emerged recently, where point clouds are no longer treated as mere compact sets but rather as empirical measures. A notion of distance to such measures has been defined and shown to be stable with respect to perturbations of the measure. This distance can easily be computed pointwise in the case of a point cloud, but its sublevel-sets, which carry the geometric information about the measure, remain hard to compute or approximate. This makes it challenging to adapt many powerful techniques based on the Euclidean distance to a point cloud to the more general setting of the distance to a measure on a metric space. We propose [28] an efficient and reliable scheme to approximate the topological structure of the family of sublevel-sets of the distance to a measure. We obtain an algorithm for approximating the persistent homology of the distance to an empirical measure that works in arbitrary metric spaces. Precise quality and complexity guarantees are given with a discussion on the behavior of our approach in practice.

7.2.3. *Topological analysis of scalar fields with outliers*

Participants: Frédéric Chazal, Steve Oudot.

In collaboration with M. Buchet, T.K. Dey, F. Fan, Y. Wang (Ohio State University).

Given a real-valued function f defined over a manifold M embedded in Euclidean space, we are interested in recovering structural information about f from the sole information of its values on a finite sample P [27]. Existing methods provide approximation to the persistence diagram of f when the noise is bounded in both the functional and geometric domains. However, they fail in the presence of aberrant values, also called outliers, both in theory and practice. We propose a new algorithm that deals with outliers. We handle aberrant functional values with a method inspired from the k -nearest neighbors regression and the local median filtering, while the geometric outliers are handled using the distance to a measure. Combined with topological results on nested filtrations, our algorithm performs robust topological analysis of scalar fields in a wider range of noise models than handled by current methods. We provide theoretical guarantees on the quality of our approximation and some experimental results illustrating its behavior.

7.2.4. *Zigzag Persistence via Reflections and Transpositions*

Participants: Clément Maria, Steve Oudot.

We introduce [33] a simple algorithm for computing zigzag persistence, designed in the same spirit as the standard persistence algorithm. Our algorithm reduces a single matrix, maintains an explicit set of chains encoding the persistent homology of the current zigzag, and updates it under simplex insertions and removals. The total worst-case running time matches the usual cubic bound.

A noticeable difference with the standard persistence algorithm is that we do not insert or remove new simplices "at the end" of the zigzag, but rather "in the middle". To do so, we use arrow reflections and transpositions, in the same spirit as reflection functors in quiver theory. Our analysis introduces a new kind of reflection called the "weak-diamond", for which we are able to predict the changes in the interval decomposition and associated compatible bases. Arrow transpositions have been studied previously in the context of standard persistent homology, and we extend the study to the context of zigzag persistence. For both types of transformations, we provide simple procedures to update the interval decomposition and associated compatible homology basis.

7.2.5. *Stable Topological Signatures for Points on 3D Shapes*

Participants: Mathieu Carrière, Steve Oudot, Maksims Ovsjanikovs.

Comparing points on 3D shapes is among the fundamental operations in shape analysis. To facilitate this task, a great number of local point signatures or descriptors have been proposed in the past decades. However, the vast majority of these descriptors concentrate on the local geometry of the shape around the point, and thus are insensitive to its connectivity structure. By contrast, several *global* signatures have been proposed that successfully capture the overall topology of the shape and thus characterize the shape as a whole. We propose [29], [43] the first point descriptor that captures the topology structure of the shape as 'seen' from a single point, in a multiscale and provably stable way. We also demonstrate how a large class of topological signatures, including ours, can be mapped to vectors, opening the door to many classical analysis and learning methods. We illustrate the performance of this approach on the problems of supervised shape labeling and shape matching. We show that our signatures provide complementary information to existing ones and allow to achieve better performance with less training data in both applications.

7.2.6. *Structure and Stability of the 1-Dimensional Mapper*

Participants: Mathieu Carrière, Steve Oudot.

Given a continuous function $f : X \rightarrow \mathbb{R}$ and a cover I of its image by intervals, the Mapper is the nerve of a refinement of the pullback cover $f^{-1}(I)$. Despite its success in applications, little is known about the structure and stability of this construction from a theoretical point of view. As a pixelized version of the Reeb graph of f , it is expected to capture a subset of its features (branches, holes), depending on how the interval cover is positioned with respect to the critical values of the function. Its stability should also depend on this positioning. We propose [44] a theoretical framework that relates the structure of the Mapper to the one of the Reeb graph, making it possible to predict which features will be present and which will be absent in the Mapper given the function and the cover, and for each feature, to quantify its degree of instability. Using this framework, we can derive guarantees on the structure of the Mapper, on its stability, and on its convergence to the Reeb graph as the granularity of the cover I goes to zero.

7.2.7. *Persistence Theory: From Quiver Representations to Data Analysis*

Participant: Steve Oudot.

Persistence theory emerged in the early 2000s as a new theory in the area of applied and computational topology. This book [35] provides a broad and modern view of the subject, including its algebraic, topological, and algorithmic aspects. It also elaborates on applications in data analysis. The level of detail of the exposition has been set so as to keep a survey style, while providing sufficient insights into the proofs so the reader can understand the mechanisms at work.

7.3. Data Structures and Robust Geometric Computation

7.3.1. *A probabilistic approach to reducing the algebraic complexity of computing Delaunay triangulations*

Participant: Jean-Daniel Boissonnat.

In collaboration with Ramsay Dyer (Johann Bernoulli Institute, University of Groningen, Netherlands) and Arijit Ghosh (Max-Planck-Institut für Informatik, Saarbrücken, Germany).

Computing Delaunay triangulations in \mathbb{R}^d involves evaluating the so-called `in_sphere` predicate that determines if a point x lies inside, on or outside the sphere circumscribing $d + 1$ points p_0, \dots, p_d . This predicate reduces to evaluating the sign of a multivariate polynomial of degree $d + 2$ in the coordinates of the points x, p_0, \dots, p_d . Despite much progress on exact geometric computing, the fact that the degree of the polynomial increases with d makes the evaluation of the sign of such a polynomial problematic except in very low dimensions. In this paper, we propose a new approach that is based on the witness complex, a weak form of the Delaunay complex introduced by Carlsson and de Silva. The witness complex $\text{Wit}(L, W)$ is defined from two sets L and W in some metric space X : a finite set of points L on which the complex is built, and a set W of witnesses that serves as an approximation of X . A fundamental result of de Silva states that $\text{Wit}(L, W) = \text{Del}(L)$ if $W = X = \mathbb{R}^d$. In [25], [41], we give conditions on L that ensure that the witness complex and the Delaunay triangulation coincide when W is a finite set, and we introduce a new perturbation scheme to compute a perturbed set L' close to L such that $\text{Del}(L') = \text{Wit}(L', W)$. Our perturbation algorithm is a geometric application of the Moser-Tardos constructive proof of the Lovász local lemma. The only numerical operations we use are (squared) distance comparisons (i.e., predicates of degree 2). The time-complexity of the algorithm is sublinear in $|W|$. Interestingly, although the algorithm does not compute any measure of simplex quality, a lower bound on the thickness of the output simplices can be guaranteed.

7.3.2. Smoothed complexity of convex hulls

Participants: Marc Glisse, Rémy Thomasse.

In collaboration with O. Devillers (VEGAS Project-team) and X. Goaoc (Université Marne-la-Vallée)

We establish an upper bound on the smoothed complexity of convex hulls in \mathbb{R}^d under uniform Euclidean (ℓ^2) noise. Specifically, let $\{p_1^*, p_2^*, \dots, p_n^*\}$ be an arbitrary set of n points in the unit ball in \mathbb{R}^d and let $p_i = p_i^* + x_i$, where x_1, x_2, \dots, x_n are chosen independently from the unit ball of radius δ . We show that the expected complexity, measured as the number of faces of all dimensions, of the convex hull of $\{p_1, p_2, \dots, p_n\}$ is $O\left(n^{2-\frac{4}{d+1}}(1+1/\delta)^{d-1}\right)$; the magnitude δ of the noise may vary with n . For $d = 2$ this bound improves to $O\left(n^{\frac{2}{3}}(1+\delta^{-\frac{2}{3}})\right)$.

We also analyze the expected complexity of the convex hull of ℓ^2 and Gaussian perturbations of a nice sample of a sphere, giving a lower-bound for the smoothed complexity. We identify the different regimes in terms of the scale, as a function of n , and show that as the magnitude of the noise increases, that complexity varies monotonically for Gaussian noise but non-monotonically for ℓ^2 noise [31], [38].

7.3.3. Realization Spaces of Arrangements of Convex Bodies

Participant: Alfredo Hubard.

In collaboration with M. Dobbins (PosTech, South Korea) and A. Holmsen (KAIST, South Korea)

In [23], we introduce combinatorial types of arrangements of convex bodies, extending order types of point sets to arrangements of convex bodies, and study their realization spaces. Our main results witness a trade-off between the combinatorial complexity of the bodies and the topological complexity of their realization space. On one hand, we show that every combinatorial type can be realized by an arrangement of convex bodies and (under mild assumptions) its realization space is contractible. On the other hand, we prove a universality theorem that says that the restriction of the realization space to arrangements of convex polygons with a bounded number of vertices can have the homotopy type of any primary semialgebraic set.

7.3.4. Limits of order types

Participant: Alfredo Hubard.

In collaboration with X. Goaoc (Institut G. Monge), R. de Joannis de Verclos (CNRS-INPG), J-S. Sereni (LORIA), and J. Volec (ETH)

The notion of limits of dense graphs was invented, among other reasons, to attack problems in extremal graph theory. It is straightforward to define limits of order types in analogy with limits of graphs, and in [24] we examine how to adapt to this setting two approaches developed to study limits of dense graphs. We first consider flag algebras, which were used to open various questions on graphs to mechanical solving via semidefinite programming. We define flag algebras of order types, and use them to obtain, via the semidefinite method, new lower bounds on the density of 5- or 6-tuples in convex position in arbitrary point sets, as well as some inequalities expressing the difficulty of sampling order types uniformly. We next consider graphons, a representation of limits of dense graphs that enable their study by continuous probabilistic or analytic methods. We investigate how planar measures fare as a candidate analogue of graphons for limits of order types. We show that the map sending a measure to its associated limit is continuous and, if restricted to uniform measures on compact convex sets, a homeomorphism. We prove, however, that this map is not surjective. Finally, we examine a limit of order types similar to classical constructions in combinatorial geometry (Erdős-Szekeres, Horton...) and show that it cannot be represented by any somewhere regular measure; we analyze this example via an analogue of Sylvester's problem on the probability that k random points are in convex position.

GRACE Project-Team

7. New Results

7.1. Weight distribution of Algebraic-Geometry codes

V. Ducet worked on the weight distribution of geometric codes following a method initiated by Duursma. More precisely he implemented his method in magma and was able to compute the weight distribution of the geometric codes coming from two optimal curves of genus 2 and 3 over the finite fields of size 16 and 9 respectively. The aim is to compute the weight distribution of the Hermitian code over the finite field of size 16, for which computational improvements of the implementation are necessary.

7.2. Faster elliptic and hyperelliptic curve cryptography

B. Smith made several contributions to the development of faster arithmetic on elliptic curves and genus 2 Jacobians in 2015. First, an extended and more detailed treatment of his \mathbb{Q} -curve construction for endomorphism-accelerated elliptic curves (previously presented at ASIACRYPT 2013, and the basis of a successful implementation with C. Costello and H. Hisil presented at EUROCRYPT 2014) appeared in the Journal of Cryptology. A simplified approach to essential precomputations was published in the proceedings of AGCT-14. Finally, with C. Costello and P.-N. Chung, he gave a new, efficient, uniform, and constant-time scalar multiplication algorithm for genus 2 Jacobians exploiting fast Kummer surface arithmetic and features of differential addition chains.

7.3. Quantum factoring

Integer factorization via Shor's algorithm is a benchmark problem for general quantum computers, but surprisingly little work has been done on optimizing the algorithm for use as a serious factoring tool once large quantum computers are built (rather than as a proof of concept). In the meantime, given the limited size of contemporary quantum computers and the practical difficulties involved in building them, any optimizations to quantum factoring algorithms can lead to significant practical improvements. In a new interdisciplinary project with physicists F. Grosshans and T. Lawson, F. Morain and B. Smith have derived a simple new quantum factoring algorithm for cryptographic integers; its expected runtime is lower than Shor's factoring algorithm, and it should also be easier to implement in practice.

7.4. Cryptanalysis of code based cryptosystems by filtration attacks

The McEliece encryption scheme based on binary Goppa codes was one of the first public-key encryption schemes [35]. Its security rests on the difficulty of decoding an arbitrary code. The original proposal uses classical Goppa codes, and while it still remains unbroken, it requires a huge size of key. On the other hand, many derivative systems based on other families of algebraic codes have been subject to key recovery attacks. Up to now, key recovery attacks were based either on a variant of Sidelnikov and Shestakov's attack [36], where the first step involves the computation of minimum-weight codewords, or on the resolution of a system of polynomial equations using Gröbner bases.

In [3], A. Couvreur, P. Gaborit, V. Gauthier, A. Otmani and J.-P. Tillich introduced a new paradigm of attack called *filtration attacks*. The general principle decomposes in two steps:

1. **Distinguishing** the public code from a random one using the square code operation.
2. **Computing a filtration** of the public code using the distinguisher, and deriving from this filtration an efficient decoding algorithm for the public code.

This new style of attack allowed A. Couvreur, A. Otmani and J.-P. Tillich to break (in polynomial time) McEliece based on wild Goppa codes over quadratic extensions [7] and more recently to break the BBCRS cryptosystem [20]. A. Couvreur, Irene Márquez–Corbella, and R. Pellikaan broke McEliece based on algebraic geometry codes from curves of arbitrary genus [5], [6] by reconstructing optimal polynomial time decoding algorithms from the raw data of a generator matrix.

7.5. Quantum LDPC codes

Quantum codes are the analogous of error correcting codes for a quantum computer. A well known family of quantum codes are the CSS codes due to Calderbank, Shor and Steane can be represented by a pair of matrices (H_X, H_Z) such that $H_X H_Z^T = 0$. As in classical coding theory, if these matrices are sparse, then the code is said to be LDPC. An open problem in quantum coding theory is to get a family of quantum LDPC codes whose asymptotic minimum distance is in $\Omega(n^\alpha)$ for some $\alpha > 1/2$. No such family is known and actually, only few known families of quantum LDPC codes have a minimum distance tending to infinity.

In an article in preparation, Benjamin Audoux (I2M, Marseille) and A. Couvreur investigate a problem suggested by Bravyi and Hastings. They studied the behaviour of iterated tensor powers of CSS codes and prove in particular that such families always have a minimum distance tending to infinity. They propose also 3 families of LDPC codes whose minimum distance is in $\Omega(n^\beta)$ for all $\beta < 1/2$.

7.6. Discrete Logarithm computations in finite fields with the NFS algorithm

The best discrete logarithm record computations in prime fields and large characteristic finite fields are obtained with Number Field Sieve algorithm (NFS) at the moment. This algorithm is made of four steps:

1. polynomial selection;
2. relation collection (with a sieving technique);
3. linear algebra (computing the kernel of a huge matrix, of millions of rows and columns);
4. individual discrete logarithm computation.

The two more time consuming steps are the relation collection step and the linear algebra step. The polynomial selection is quite fast but is very important since it determines the complexity of the algorithm. Selecting better polynomials is a key to improve the overall running-time of the NFS algorithm. The final step: individual discrete logarithm, was though to be quite fast but F. Morain and A. Guillevic showed that it has an increasing complexity with respect to the extension degree of the finite field. A. Guillevic proposed a new method to reduce considerably the complexity, with at most a factor two speed-up in the exponent [22].

In 2015, F. Morain and A. Guillevic released with P. Gaudry and R. Barbulescu a major discrete logarithm record in a quadratic finite field $\text{GF}(p^2)$ of 180 decimal digits (dd), corresponding to 595 bits. This was presented at the international conference Eurocrypt [19].

7.6.1. DL Record computation in a quadratic finite field $\text{GF}(p^2)$

In order to compare the practical running time of discrete logarithm computation in prime fields and quadratic finite fields, F. Morain and A. Guillevic with P. Gaudry and R. Barbulescu launched a DL record in a 180dd finite field. The last DL record in a prime field was held by the CAMEL team of Nancy, in 2014, in a 180 dd prime field. The parameters chosen for the quadratic finite field are the following.

$$\begin{aligned}
 p &= 314159265358979323846264338327950288419716939937510582097494459230 \setminus \\
 &\quad 781640628620899877709223 \\
 \ell &= 392699081698724154807830422909937860524646174921888227621868074038 \setminus \\
 &\quad 47705078577612484713653 \\
 p-1 &= 6 \cdot h_0 \text{ with } h_0 \text{ a } 89 \text{ dd prime number} \\
 p+1 &= 8 \cdot \ell
 \end{aligned}$$

The discrete logarithm computation was made modulo ℓ , the largest prime factor of the multiplicative subgroup $GF(p^2)^*$, so that a DL computation with generic methods of complexity $O(\sqrt{\ell})$ was impracticable.

The two polynomials used in the NFS algorithm were chosen to be the following:

$$\begin{aligned} f &= x^4 + 1 \\ g &= 448225077249286433565160965828828303618362474 x^2 - 29606109908476368046927513730655796265782 \\ &\quad + 448225077249286433565160965828828303618362474 . \end{aligned}$$

We indeed designed a new polynomial selection method, that we called the Conjugation method. It is very well suited for quadratic and cubic finite fields $GF(p^2)$ and $GF(p^3)$ for the size range of the records.

We finally computed the discrete logarithm in basis $G = T + 2$ of the target $s = \lfloor (\pi(2^{298})/8) \rfloor t + \lfloor (\gamma \cdot 2^{298}) \rfloor$

$$\begin{aligned} \log_G s &\equiv 2762142436179128043003373492683066054037581738194144186101 \setminus \\ &\quad 9832278568318885392430499058012 \pmod{\ell}. \end{aligned}$$

The running time was very surprising: our record was much faster than the concurrent DL computation in a prime field of the same global size of 180dd, and even faster than the RSA modulus factorization of the same size.

Table 2. Comparison of running time for integer factorization (NFS-IF), discrete logarithm in prime field (NFS-DL(p)) and in quadratic field (NFS-DL(p²)) of same global size 180 dd.

Algorithm	relation collection	linear algebra	total
NFS-IF	5 years	5.5 months	5.5 years
NFS-DL(p)	50 years	80 years	130 years
NFS-DL(p ²)	157 days	18 days (GPU)	0.5 years

7.6.2. Individual discrete logarithm computation

A big difference between prime fields and finite fields of small extension such as $GF(p^3)$, $GF(p^4)$ and $GF(p^6)$ is the complexity of the final step of the NFS algorithm: computing the individual discrete logarithm of the target, given the large table of discrete logarithm of *small* elements. This table was obtained at the end of the linear algebra step. The target needs to be decomposed into small enough elements whose discrete logarithm is in the table, so that one can recompose the discrete logarithm of the target. This decomposition is quite fast for prime fields but we realized that it becomes more and more time consuming when the extension degree increase. A. Guillevic developed a new technique to improve considerably this step. The main idea is to use the structure of the finite field: the subfields. These improvements were presented at the Asiacrypt 2015 conference in Auckland, New Zealand and published in the proceedings [22].

7.7. Information sets of multiplicity codes

The codes we used in our PIR protocols, namely Reed-Muller and their generalization Multiplicity codes, are locally *correctable*: that means that local decoding allows to retrieve encoded symbols. In most applications, it is very desirable to retrieve *information* symbols. Another line of work in this topic was thus to find an encoding method for multiplicity codes so as to directly recover an information symbol from local correction, and not an encoded symbol. To do so we defined information sets for multiplicity codes, and design a systematic encoding based on this information set. This work was presented at ISIT'2015 in Hong-Kong in June [18].

7.8. Rank metric codes over infinite fields

Rank metric and Gabidulin codes over the rationals promise interesting applications to space-time coding. We have constructed optimal codes, similar to Gabidulin codes, in the case of infinite fields. We use algebraic extensions, and we have determined the condition on the considered extension to enable this construction. For example: we can design codes with complex coefficients, using number fields and Galois automorphisms. Then, in the rank metric setting, codewords can be seen as matrices. In this setting, a channel introduces errors (a matrix of small rank r added to the codeword) and erasures (s_r rows and s_c columns of the matrix are erased). We have developed an algorithm (adapted from the Welch–Berlekamp algorithm) to recover the right codeword in the presence of an error of rank weight up to $r + s_c + s_r \leq d - 1$, where d is the minimal distance of the code. As opposed to the finite field case, we are confronted by coefficient size growth. We solve this problem by computing modulo prime ideals. Using these codes we can completely bypass intermediate constructions using finite fields, which were the stumbling-block in classic constructions.

We also have used this framework to build rank-metric codes over the field of rational functions, using algebraic function fields with cyclic Galois group (Kummer and Artin extensions). These codes can be seen as a generator of infinitely many convolutional codes.

7.9. Hash function cryptanalysis

Cryptographic hash functions are versatile primitives that are used in many cryptographic protocols. The security of a hash function h is usually evaluated through two main notions: its preimage resistance (given a target t , the difficulty of finding a message m s.t. $h(m) = t$) and its collision resistance (the difficulty of finding two messages m, m' s.t. $h(m) = h(m')$).

A popular hash function is the SHA-1 algorithm. Although theoretical collision attacks were found in 2005, it is still being used in some applications, for instance as the hash function in some TLS certificates. Hence cryptanalysis of SHA-1 is still a major topic in cryptography.

In 2015, we improved the state-of-the-art on SHA-1 analysis in two ways:

- T. Espitau, P.-A. Fouque and P. Karpman improved the previous preimage attacks on SHA-1, reaching up to 62 rounds (out of 80), up from 57. The corresponding paper was published at CRYPTO 2015 [21].
- P. Karpman, T. Peyrin and M. Stevens developed collision attacks on the compression function of SHA-1 (i.e. freestart collisions). This exploits a model that is slightly more generous to the attacker in order to find explicit collisions on more rounds than what was previously possible. A first work resulted in freestart collisions for SHA-1 reduced to 76 steps; this attack takes less than a week to compute on a common GPU. The corresponding paper was published at CRYPTO 2015 [24]. This was later improved to attack the full compression function. Although the attack is more expensive it is still practical, taking less than two weeks on a 64 GPU cluster. The corresponding paper is currently under review for EUROCRYPT 2016 [32].

7.10. Block cipher design and analysis

Block ciphers are one of the most basic cryptographic primitives, yet block cipher analysis is still a major research topic. In recent years, the community also shifted focus to the more general setting of *authenticated encryption*, where one specifies an (set of) algorithm(s) providing both encryption and authentication for messages of arbitrary length. A major current event in that direction is the CAESAR academic competition, which aims to select a portfolio of good algorithms.

During this year, we helped to improve the state of the art in block cipher research in several ways:

- P. Karpman found a very efficient related-key attack on the CAESAR candidate Prøst-OTR. A related-key model is very generous to the attacker, but the attack in this case can be run instantaneously. The corresponding paper was published at ISC 2015 [23]

- B. Minaud, P. Derbez, P.-A. Fouque and P. Karpman developed a family of attacks that breaks all the remaining unbroken instances of the ASASA construction, that was presented at ASIACRYPT 2014. Using algebraic properties of the ciphers, for each type of instance, the attack allows to recover an algorithm equivalent to the secret key in near-practical time. This applies to a multivariate public-key scheme, a classical block cipher and small block ciphers used in white-box constructions. The corresponding paper was published at ASIACRYPT 2015 and was honoured as one of the three best papers of the conference [25].
- P. Karpman developed a compact 8-bit S-box with branch number three, which can be used as a basis to construct a lightweight block cipher particularly efficient on 8-bit microcontrollers. The corresponding paper is currently under review for FSE 2016.

HYCOMES Team

6. New Results

6.1. Embedded Systems Design

6.1.1. *Loosely Time-Triggered Architectures: Improvements and Comparisons*

Participant: Albert Benveniste.

Loosely Time-Triggered Architectures (LTTAs) are a proposal for constructing distributed embedded control systems. They build on the quasi-periodic architecture, where computing units execute 'almost periodically', by adding a thin layer of middleware that facilitates the implementation of synchronous applications. In [7], we have shown how the deployment of a synchronous application on a quasi-periodic architecture can be modeled using a synchronous formalism. Then we have detailed two protocols, Back-Pressure LTTA, reminiscent of elastic circuits, and Time-Based LTTA, based on waiting. Compared to previous work, we presented controller models that can be compiled for execution and a simplified version of the Time-Based protocol. We also compared the LTTA approach with architectures based on clock synchronization.

6.2. Hybrid Systems Modeling

Participants: Ayman Aljarbough, Albert Benveniste, Benoît Caillaud, Khalil Ghorbal.

6.2.1. *Robust Simulation for Hybrid Systems: Chattering Path Avoidance*

The sliding mode approach is recognized as an efficient tool for treating the chattering behavior in hybrid systems. However, the amplitude of chattering, by its nature, is proportional to magnitude of discontinuous control. A possible scenario is that the solution trajectories may successively enter and exit as well as slide on switching manifolds of different dimensions. Naturally, this arises in dynamical systems and control applications whenever there are multiple discontinuous control variables. The main contribution of [9] is to provide a robust computational framework for the most general way to extend a flow map on the intersection of p intersected $(n-1)$ -dimensional switching manifolds in at least p dimensions. We explored a new formulation to which we can define unique solutions for such particular behavior in hybrid systems and investigate its efficient computation/simulation. An extended version of this work has been presented at the Baltic Young Scientists Conference [8].

6.2.2. *A Hierarchy of Proof Rules for Checking Positive Invariance of Algebraic and Semi-Algebraic Sets*

In [6], we studied sound proof rules for checking positive invariance of algebraic and semi-algebraic sets, that is, sets satisfying polynomial equalities and those satisfying finite boolean combinations of polynomial equalities and inequalities, under the flow of polynomial ordinary differential equations. Problems of this nature arise in formal verification of continuous and hybrid dynamical systems, where there is an increasing need for methods to expedite formal proofs. We study the trade-off between proof rule generality and practical performance and evaluate our theoretical observations on a set of benchmarks. The relationship between increased deductive power and running time performance of the proof rules is far from obvious; we discuss and illustrate certain classes of problems where this relationship is interesting.

6.2.3. A Formally Verified Hybrid System for Safe Advisories in the Next-Generation Airborne Collision Avoidance System

The Next-Generation Airborne Collision Avoidance System (ACAS X) is intended to be installed on all large aircraft to give advice to pilots and prevent mid-air collisions with other aircraft. It is currently being developed by the Federal Aviation Administration (FAA). In [16] we determined the geometric configurations under which the advice given by ACAS X is safe under a precise set of assumptions and formally verify these configurations using hybrid systems theorem proving techniques. We considered subsequent advisories and showed how to adapt our formal verification to take them into account. We examined the current version of the real ACAS X system and discussed some cases where our safety theorem conflicts with the actual advisory given by that version, demonstrating how formal, hybrid systems proving approaches are helping to ensure the safety of ACAS X. Our approach is general and could also be used to identify unsafe advice issued by other collision avoidance systems or confirm their safety.

6.2.4. Domain Globalization: Using Languages to Support Technical and Social Coordination

When a project is realized in a globalized environment, multiple stakeholders from different organizations work on the same system. Depending on the stakeholders and their organizations, various (possibly overlapping) concerns are raised in the development of the system. In this context a Domain Specific Language (DSL) supports the work of a group of stakeholders who are responsible for addressing a specific set of concerns. We contributed to a book chapter [11], identifying the open challenges arising from the coordination of globalized domain-specific languages. We identified two types of coordination: technical coordination and social coordination. After presenting an overview of the current state of the art, we discussed first the open challenges arising from the composition of multiple DSLs, and then the open challenges associated to the collaboration in a globalized environment.

6.3. Contracts for Systems Design

Participants: Albert Benveniste, Benoît Caillaud.

6.3.1. Contracts for Systems Design: Theory, Methodology and Application Cases

Aircrafts, trains, cars, plants, distributed telecommunication military or health care systems, and more, involve systems design as a critical step. Complexity has caused system design times and costs to go severely over budget so as to threaten the health of entire industrial sectors. Heuristic methods and standard practices do not seem to scale with complexity so that novel design methods and tools based on a strong theoretical foundation are sorely needed. Model-based design as well as other methodologies such as layered and compositional design have been used recently but a unified intellectual framework with a complete design flow supported by formal tools is still lacking. Recently an “orthogonal” approach has been proposed that can be applied to all methodologies introduced thus far to provide a rigorous scaffolding for verification, analysis and abstraction/refinement: contract-based design. Several results have been obtained in this domain but a unified treatment of the topic that can help in putting contract-based design in perspective is missing. We have published two research reports [13], [12], that intend to provide such treatment where contracts are precisely defined and characterized so that they can be used in design methodologies such as the ones mentioned above with no ambiguity. In addition, the first report [13] provides an important link between interface and contract theories to show similarities and correspondences. This report is complemented by a companion report [12] where contract based design is illustrated through use cases.

6.3.2. Contracts for Schedulability Analysis

In [10] we proposed a framework of Assume / Guarantee contracts for schedulability analysis. Unlike previous work addressing compositional scheduling analysis, our objective is to provide support for the OEM / supplier subcontracting relation. The adaptation of Assume / Guarantee contracts to schedulability analysis requires some care, due to the handling of conflicts caused by shared resources. We illustrate our framework in the context of Autosar methodology now popular in the automotive industry sector.

LFANT Project-Team

6. New Results

6.1. Class groups and other invariants of number fields

Participants: Karim Belabas, Jean-Paul Cerri, Henri Cohen, Pinar Kılıçer, Pierre Lezowski.

Ohno and Nakagawa have proved relations between the counting functions of certain cubic fields. These relations may be viewed as complements to the Scholz reflection principle, and Ohno and Nakagawa deduced them as consequences of 'extra functional equations' involving the Shintani zeta functions associated to the prehomogeneous vector space of binary cubic forms. The paper [14] by Henri Cohen, Simon Rubinstein-Salzedo and Frank Thorne proves an identity relating certain degree fields with Galois groups D and F , respectively, for any odd prime, giving another proof of the Ohno–Nakagawa relation between the counting functions of certain cubic fields.

Pinar Kılıçer and Marco Streng have solved a variant of the class number 1 problem for quartic CM fields with a geometric motivation [27]; the question is whether a certain class group is trivial, which corresponds to a genus 2 curve with that complex multiplication being defined over a real-quadratic number field (instead of an extension). Using classical techniques provides a bound on the discriminant of such fields, which they refine taking ramification into account to obtain a practically useful bound. A carefully crafted enumeration algorithm finishes the proof.

In the article [28], P. Lezowski studies the Euclidean properties of matrix algebras $M_n(R)$ over commutative rings R . In particular, he shows that for any integer $n > 1$, $M_n(R)$ is a left and right Euclidean ring if and only if R is principal. The proof is constructive and allows to better understand the Euclidean order types of matrix algebras. Similar ideas are also applied to prove k -stage Euclidean properties of $M_n(R)$, linking them with Bézout property for the ring R . The article [28] has been submitted to *Journal of Algebra*.

The article by Aurel Page on the computation of arithmetic Kleinian groups has appeared [21].

6.2. Complex L -functions and certified arithmetic

Participants: Bill Allombert, Karim Belabas, Henri Cohen, Fredrik Johansson.

Fredrik Johansson's paper [23] has been published and presented at the 22nd IEEE Symposium on Computer Arithmetic (ARITH22), Lyon, France. This paper describes a new implementation of the elementary transcendental functions \exp , \sin , \cos , \log and atan for variable precision up to approximately 4096 bits. Compared to the MPFR library, it achieves a maximum speedup ranging from a factor 3 for \cos to 30 for atan .

Bill Allombert, Karim Belabas, Henri Cohen and Pascal Molin (Paris 7) have implemented a new framework in PARI/GP to compute and manipulate complex L -functions and the associated ϑ and Λ functions, exporting 25 new functions to the GP computer algebra system.

6.3. Elliptic curve and Abelian varieties cryptology

Participants: Jean-Marc Couveignes, Andreas Enge, Enea Milio, Damien Robert.

In [29] David Lubicz and Damien Robert explain how to improve the arithmetic of Abelian and Kummer varieties. The speed of the arithmetic is a crucial factor in the performance of cryptosystems based on abelian varieties. Depending on the cryptographic application, the speed record holders are elliptic curves (in the Edwards model) or the Kummer surface of an hyperelliptic curves of genus 2 (in the level 2 theta model). One drawback of the Kummer surface is that only scalar multiplications are available, which may be a problem in certain cryptographic protocols. The previous known models to work on the Jacobian rather than the Kummer surface (Mumford coordinates or the theta model of level 4) are too slow and not competitive with elliptic curves. This paper explains how to use geometric properties (like projective normality) to speed up the arithmetic. In particular it introduces a novel addition algorithm on Kummer varieties (compatible addition), and uses it to speed up multi-exponentiations in Kummer varieties and to obtain new models of abelian surfaces in which the scalar multiplication is as fast as on the Kummer surface. This paper was written last year but heavily revised in 2015 and has been accepted (up to minor revisions) in the journal *Finite Fields and Their Applications*.

The paper [19] by David Lubicz and Damien Robert about computing certain isogenies in quasi optimal time has been published in the *LMS Journal of Computation and Mathematics* and the paper [18] by the same authors about optimal pairing computation on abelian varieties has been published in the *Journal of Symbolic Computation*. This paper expands the article [15] by Romain Cosset and Damien Robert about the computation of (ℓ, ℓ) -isogenies in dimension 2 which has been published in *Mathematics of Computation*.

Enea Milio has published one of the main results of his PhD thesis [20]. He has generalised the work of Régis Dupont for computing modular polynomials in dimension 2 to new invariants. He describes an algorithm to compute modular polynomials for invariants derived from theta constants and proves under some heuristics that this algorithm is quasi-linear in its output size. Some properties of the modular polynomials defined from quotients of theta constants are analysed and experiments with an implementation are related.

The paper [16] by Jean-Marc Couveignes and Tony Ezome explaining how to efficiently evaluate functions, including Weil functions and canonical theta functions, on Jacobian varieties and their quotients has been published in the *LMS Journal of Computation and Mathematics*. This paper also describes a quasi-optimal algorithm to compute (l, l) -isogenies between Jacobians of genus two curves, using a compact representation and differential characterisation of isogenies.

In [26], Sorina Ionica and Emmanuel Thomé look at the structure of isogeny graphs of genus 2 Jacobians with maximal real multiplication. They generalise a result of Kohel's describing the structure of the endomorphism rings of the isogeny graph of elliptic curves. Their setting considers genus 2 jacobians with complex multiplication, with the assumptions that the real multiplication subring is maximal and has class number 1. Over finite fields, they derive a depth first search algorithm for computing endomorphism rings locally at prime numbers, if the real multiplication is maximal.

6.4. Cryptology with quadratic fields

Participant: Guilhem Castagnos.

In [22] Guilhem Castagnos and Fabien Laguillaumie design a linearly homomorphic encryption scheme the security of which relies on the hardness of the decisional Diffie-Hellman problem. The approach requires some special features of the underlying group. In particular, its order is unknown and it contains a subgroup in which the discrete logarithm problem is tractable. Therefore, their instantiation holds in the class group of a non-maximal order of an imaginary quadratic field. Its algebraic structure makes it possible to obtain such a linearly homomorphic scheme in which the message space is the whole set of integers modulo a prime p and which supports an unbounded number of additions modulo p from the ciphertexts. A notable difference with previous work is that, for the first time, the security does not depend on the hardness of the factorisation of integers. As a consequence, under some conditions, the prime p can be scaled to fit the application needs. This paper has been presented at the cryptographer's track at the RSA Conference 2015.

MARELLE Project-Team

6. New Results

6.1. IDE for Coq

Participants: Enrico Tassi, Alexander Faithfull [ITU Copenhagen], Jesper Bengtson [ITU Copenhagen], Carst Tankink.

User interfaces for interactive proof assistants should rely on the advanced software available in integrated development environments. We collaborated with researchers from Copenhagen to build an Eclipse-based environment for the Coq system. This exploits the quick compilation chain that was developed for Coq 8.5. This work has been published in [15].

6.2. ELPI, Fast, Embeddable, λ -Prolog Interpreter

Participants: Enrico Tassi, Cvetan Dunchev [University of Bologna], Ferruccio Guidi [University of Bologna], Claudio Sacerdoti Coen [University of Bologna].

We developed a new interpreter that runs consistently faster than the other available implementations of λ -prolog. The key insight is the identification of a fragment of the language, which we call reduction-free fragment, that occurs quite naturally and that admits constant time reduction and unification rules. In the long run, we hope that this will contribute to developing elaborators that support a more efficient and adaptable usage of interactive proof tools. This work is published in [14].

6.3. Verified Proofs of Higher-Order Masking

Participants: Gilles Barthe [IMDEA Software, Madrid], Sonia Belaïd [Thales Communication], François Dupressoir [IMDEA Software, Madrid], Pierre-Alain Fouque [Université de Rennes, IUF], Benjamin Grégoire, Pierre-Yves Strub [IMDEA Software, Madrid].

We study the problem of automatically verifying higher-order masking countermeasures. We propose a method based on program verification techniques, to check the independence of sets of intermediate variables from secrets. This new language-based technique makes it possible to implement several algorithms that reduce the number of sets of variables that need consideration. The tool also has the capability to give useful information when proofs fail, for instance by discovering possible attacks. This is based on EasyCrypt. This work has been published in [8].

6.4. Relational Reasoning via Probabilistic Coupling

Participants: Gilles Barthe [IMDEA Software, Madrid], Thomas Espitau [ENS Cachan], Benjamin Grégoire, Justin Hsu [University of Pennsylvania], Léo Stefanescu [ENS Lyon], Pierre-Yves Strub [IMDEA Software, Madrid].

Probabilistic coupling is a powerful tool for analyzing pairs of probabilistic processes. While the mathematical definition of coupling looks rather complex and cumbersome to manipulate, we show that the relational program logic pRHL—the logic underlying the EasyCrypt cryptographic proof assistant—already internalizes a generalization of probabilistic coupling. With this insight, constructing couplings is no harder than constructing logical proofs. We demonstrate how to express and verify classic examples of couplings in pRHL, and we mechanically verify several couplings in EasyCrypt. This work is described in [9].

6.5. Automated Proofs of Pairing-Based Cryptography

Participants: Gilles Barthe [IMDEA Software, Madrid], Benjamin Grégoire, Benedikt Schmidt [IMDEA Software, Madrid].

We implement a new tool, called AutoG&P, which supports extremely compact, and often fully automated, proofs of cryptographic constructions based on (bilinear or multilinear) Diffie-Hellman assumptions. For instance, we provide a 100-line proof of Waters' Dual System Encryption (CRYPTO'09), and fully automatic proofs of Boneh-Boyen Identity-Based Encryption (CRYPTO'04). Finally, we provide an automated tool that generates independently verifiable EasyCrypt proofs from AutoG&P proofs. This work has been published in [10].

6.6. Improvements on CBC MAC formalized in EasyCrypt

Participants: Benjamin Grégoire, Cécile Baritel-Ruet, Pierre-Alain Fouque.

In a paper of 2003, J. Black and P. Rogaway propose variations of cipher block chaining message authentication codes for the efficient authentication of arbitrary length messages. We formalize their work in EasyCrypt, resulting in formal proofs for CBC-MAC, EMAC, ECBC, FCBC and the most efficient of these variations, XCBC.

This work required the development of new EasyCrypt theories. A small flaw in the original paper was found and a fix has been proposed. This work was partially funded by the Brutus ANR project.

6.7. Buchberger's algorithm and advanced formalization of multinomials

Participant: Laurent Théry.

We studied how the Mathematical Components library could improve the formalization of algorithms based on multivariate polynomials. In particular, we re-used Pierre-Yves Strub library of multivariate polynomials and re-did the proofs of correctness for Buchberger's algorithm. This new piece of formalized algorithm is now available at the following address <https://github.com/theyr/grobner>.

6.8. Proofs that e and π and transcendental

Participants: Sophie Bernard, Laurence Rideau, Yves Bertot, Pierre-Yves Strub [IMDEA Software, Madrid].

In the previous year, we developed formally verified proofs that e and π are transcendental. This year we cleaned up these proofs to obtain a common lemma that applies in both cases with simple hypotheses. In parallel, P.-Y. Strub streamlined the library on multivariate polynomials which plays a significant role in the case of π . This work has been published in [11].

In the future, we will probably extend this work to more general proofs of transcendence.

6.9. Algorithms for Real Algebraic Geometry

Participant: Cyril Cohen.

We formalized an efficient algorithm to count roots of a polynomial satisfying polynomial inequalities. This work was presented at the Types workshop in May and the Workshop on Algebra, Geometry, and Proofs in Symbolic computation.

6.10. Nominal sets in Coq

Participants: Cyril Cohen, Nicolas Tabareau, Matthieu Sozeau, Gabriel Lewertowski.

Cyril Cohen collaborated with members of the team $\pi.r^2$ on the implementation of nominal sets in Coq.

6.11. Formal Description of Dynamic Logic

Participants: Yves Bertot, Cyril Cohen, Jean-Yves Franceschi.

We developed a formal description of the language of dynamic logic in the Coq system.

6.12. Cubical Type Theory

Participants: Cyril Cohen, Thierry Coquand, Simon Huber, Anders Mörtberg.

We participate to the development of a software prototype, `cubicaltt`, <https://github.com/mortberg/cubicaltt>, that is expected to support an extension of type theory suited for homotopy type theory.

6.13. Finite set and finite maps

Participant: Cyril Cohen.

We extend the Math-Components library with a module concerning finite sets (in potentially infinite types) and finite maps. This module will play a crucial role in other experiments, like the experiments on dynamic logic, nominal sets, and cubical sets.

6.14. Formalization of a Newton Series Representation of Polynomials

Participants: Boris Djalal, Cyril Cohen.

We formalize an algorithm to change the representation of a polynomial to a Newton power series. This provides a way to compute efficiently polynomials whose roots are the sums or products of roots of other polynomials, and hence provides a base component of efficient computation for algebraic numbers. In order to achieve this, we formalize a notion of truncated power series and develop an abstract theory of poles of fractions. This work is described in [13].

6.15. Formal description of catalan numbers

Participant: José Grimm.

Catalan number can be defined by a recurrence, or by explicit formulas using binomial numbers. An important property is $C_{n+1} = \sum_{k \leq n} C_k C_{n-k}$. The easiest way to prove this formula is to use Dyck paths.

A Dyck path of size $2n$ is a sequence l of integers $+1$ and -1 so that the sum s_k of the k first terms is ≥ 0 for $k \leq 2n$ and $s_{2n} = 0$. The relation between Dyck paths and Catalan numbers is easy to prove and then properties of Dyck paths are quite simple to state and verify.

The proofs have been done with the Math-Components library.

6.16. Latex to XML translator

Participant: José Grimm.

This year, we released version 2.15.4 of Tralics, our LaTeX to XML translator. Array handling has been redesigned: for instance, an array preamble of the form $\{>\{\}\}c\{<\{\}\}$ is now correctly interpreted; there is a possibility to add an attribute pair to any table, row or cell; for math environments like “align”, one label and one tag per row is allowed. Tralics is also able to read an XML file, and there are some primitives for inserting the result (or part of it) into the XML code under construction.

MEXICO Project-Team

7. New Results

7.1. Highlights

Please note that three of our most important and novel results are given in the 'Highlights' section above.

7.2. Specifying and Verifying Concurrent C Programs with TLA+

Verifying software systems automatically from their source code rather than modelling them in a dedicated language gives more confidence in establishing their properties. In [37] we propose a formal specification and verification approach for concurrent C programs directly based on the semantics of C. We define a set of translation rules and implement it in a tool (C2TLA+) that automatically translates C code into a TLA+ specification. The TLC model checker can use this specification to generate a model, allowing to check the absence of runtime errors and dead code in the C program in a given configuration. In addition, we show how translated specifications interact with manually written ones to: check the C code against safety or liveness properties; provide concurrency primitives or model hardware that cannot be expressed in C; and use abstract versions of translated C functions to address the state explosion problem. All these verifications have been conducted on an industrial case study, which is a part of the microkernel of the PharOS real-time system.

7.3. Active Diagnosis with Observable Quiescence

Active diagnosis of a discrete-event system consists in controlling the system such that faults can be detected. In [27] we extend the framework of active diagnosis presented in [7] by introducing modalities for actions and states and a new capability for the controller, namely observing that the system is quiescent. We design a game-based construction for both the decision and the synthesis problems that is computationally optimal. Furthermore we prove that the size and the delay provided by the active diagnoser (when it exists) are almost optimal.

7.4. Test Case Generation for Concurrent Systems Using Event Structures

In [23] we deal with the test-case generation problem for concurrent systems that are specified by true-concurrency models such as Petri nets. We show that using true-concurrency models reduces both the size and the number of test cases needed for achieving certain coverage criteria. We present a test-case generation algorithm based on Petri net unfoldings and a SAT encoding for solving controllability problems in test cases. Finally, we evaluate our algorithm against traditional test-case generation methods under interleaving semantics.

7.5. State Space Reduction Strategie for Model Checking Concurrent C Programs

Model checking is an effective technique for uncovering subtle errors in concurrent systems. Unfortunately, the state space explosion is the main bottleneck in model checking tools. In [31] we propose a state space reduction technique for model checking concurrent programs written in C. The reduction technique consists in an analysis phase, which defines an approximate agglomeration predicate. This latter states whether a statement can be agglomerated or not. We implement this predicate using a syntactic analysis, as well as a semantic analysis based on abstract interpretation. We show the usefulness of using agglomeration technique to reduce the state space, as well as to generate an abstract TLA+ specification from a C program.

7.6. Simple Priced Timed Games Are Not That Simple

Priced timed games are two-player zero-sum games played on priced timed automata (whose locations and transitions are labeled by weights modeling the costs of spending time in a state and executing an action, respectively). The goals of the players are to minimise and maximise the cost to reach a target location, respectively. In [25] we consider priced timed games with one clock and arbitrary (positive and negative) weights and show that, for an important subclass of theirs (the so-called simple priced timed games), one can compute, in exponential time, the optimal values that the players can achieve, with their associated optimal strategies. As side results, we also show that one-clock priced timed games are determined and that we can use our result on simple priced timed games to solve the more general class of so-called reset-acyclic priced timed games (with arbitrary weights and one-clock).

7.7. A Hybrid-Dynamical Model for Passenger-flow in Transportation Systems

In a network with different transportation modes, or multimodal public transportation system (MPTS), modes are linked among one another not by resources or infrastructure elements—which are not shared, e.g., between different metro lines—but by the flow of passengers between them. Now, the movements of passengers are steered by the destinations that individual passengers have, and by which they can be grouped into trip profiles. To use the strength of fluid dynamics, introduce in [30] a multiphase hybrid Petri net model, in which the vehicle dynamics is rendered by individual tokens moving in an infrastructure net, while passenger quantities are given as vectors—whose components correspond to trip profiles—and evolve at stations according to fluid dynamics. This model is intended as a building block for obtaining supervisory control, via transport operator actions, to mitigate congestion.

7.8. An Algebraic View of Space/Belief and Extrusion/Utterance for Concurrency/Epistemic Logic

In [29] we enrich spatial constraint systems with operators to specify information and processes moving from a space to another. We shall refer to these new structures as spatial constraint systems with extrusion. We shall investigate the properties of this new family of constraint systems and illustrate their applications. From a computational point of view the new operators provide for process/information extrusion, a central concept in formalisms for mobile communication. From an epistemic point of view extrusion corresponds to a notion we shall call utterance; a piece of information that an agent communicates to others but that may be inconsistent with the agent's beliefs. Utterances can then be used to express instances of epistemic notions, which are common place in social media, such as hoaxes or intentional lies. Spatial constraint systems with extrusion can be seen as complete Heyting algebras equipped with maps to account for spatial and epistemic specifications.

7.9. Preserving Partial Order Runs in Parametric Time Petri Nets

Parameter synthesis for timed systems aims at deriving parameter valuations satisfying a given property. In [22] we target concurrent systems; it is well known that concurrency is a source of state-space explosion, and partial order techniques were defined to cope with this problem. Here we use partial order semantics for parametric time Petri nets as a way to significantly enhance the result of an existing synthesis algorithm. Given a reference parameter valuation, our approach synthesizes other valuations preserving, up to interleaving, the behavior of the reference parameter valuation. We show the applicability of our approach using acyclic asynchronous circuits.

7.10. Non-Atomic Transition Firing in Contextual Nets

The firing rule for Petri nets assumes instantaneous and simultaneous consumption and creation of tokens. In the context of ordinary Petri nets, this poses no particular problem because of the system's asynchronicity, even if token creation occurs later than token consumption in the firing. With read arcs, the situation changes, and several different choices of semantics are possible. The step semantics introduced by Janicki and Koutny

can be seen as imposing a two-phase firing scheme: first, the presence of the required tokens is checked, then consumption and production of tokens happens. Pursuing this approach further, we develop in [28] a more general framework based on explicitly splitting the phases of firing, allowing to synthesize coherent steps. This turns out to define a more general non-atomic semantics, which has important potential for safety as it allows to detect errors that were missed by the previous semantics. Then we study the characterization of partial-order processes feasible under one or the other semantics.

MUTANT Project-Team

7. New Results

7.1. Weakly-Supervised Discriminative Model for Audio-to-Score Alignment

We consider a new discriminative approach to the problems of segmentation and of audio-to-score alignment. For each musical event, templates have to be built or learnt before performing any alignment. Because annotating a large database music files would be a tedious task, we develop an original approach to learn templates without annotations, but only the knowledge of the music scores associated to music files. We consider the two distinct informations provided by the music scores: (i) an exact ordered list of musical events and (ii) an approximate prior information about relative duration of events. We extend the celebrated Dynamic Time Warping algorithm (DTW) to a convex problem that learns optimal classifiers for all events while jointly aligning files, using this weak supervision only. We show that the relative duration between events can be easily used as a penalization of our cost function and allows us to drastically improve performances of our approach. We describe in details our approach and preliminary results obtained on a large-scale database in [18].

This work was done in collaboration with the SIERRA project-team at Inria Paris.

7.2. Semi-Markov Models for Real-time MIDI-to-Score Alignment

We develop a new stochastic model of symbolic (MIDI) performance of polyphonic scores, based on Semi-Markov models, to align MIDI performances of music scores. In our approach, the evolution of the music performer and the production of performed notes are modeled with a hierarchical extension of hidden semi-Markov models (HSMM). By comparing with a previously studied model based on hidden Markov model (HMM), we give theoretical reasons why the present model is advantageous to deal with complex music event such as trills, tremolos, arpeggios, and other ornaments. This is also confirmed empirically by comparing the accuracy of score following and analysing the errors. We also develop a hybrid of this HSMM-based model and the HMM-based model which is computationally more efficient and retains the advantages of the former model. The present model yields one of the state-of-the-art score following algorithms for symbolic performance and can possibly be applicable for other music recognition problems. Details and results are published in [19].

This work was done in collaboration with Eita Nakamura from the National Institute of Informatics of Tokyo, Japan.

7.3. Real-time Audio-to-Score Alignment of Singing Voice

Singing voice is specific in music: a vocal performance conveys both music (melody/pitch) and lyrics (text/phoneme) content. We develop an original approach that aims at exploiting the advantages of melody and lyric information for real-time audio-to-score alignment of singing voice. First, lyrics are added as a separate observation stream into a template-based hidden semi-Markov model (HSMM), whose observation model is based on the construction of vowel templates. Second, early and late fusion of melody and lyric information are processed during real-time audio-to-score alignment. An experiment conducted with two professional singers (male/female) shows that the performance of a lyrics-based system is comparable to that of melody-based score following systems. Furthermore, late fusion of melody and lyric information substantially improves the alignment performance. Finally, maximum a posteriori adaptation (MAP) of the vowel templates from one singer to the other suggests that lyric information can be efficiently used for any singer. Preliminary results are published in [15].

7.4. Online Methods for Audio Segmentation and Clustering

Audio segmentation is an essential problem in many audio signal processing tasks, which tries to segment an audio signal into homogeneous chunks. Rather than separately finding change points and computing similarities between segments, we focus on joint segmentation and clustering, using the framework of hidden Markov and semi-Markov models. We introduced a new incremental EM algorithm for hidden Markov models (HMMs) and showed that it compares favorably to existing online EM algorithms for HMMs. Early experimental results on musical note segmentation and environmental sound clustering are promising and will be pursued further in 2015.

Theoretical results were published in [11] in collaboration with the SIERRA project-team, and experimental results were further extended in [32]. Early experimental setups show that our algorithms outperform state-of-the-art supervised methods for Percussion Sound classification. In collaboration with IRCyNN (Nantes) we are currently studying algorithmic extensions to complex environmental sounds.

7.5. Adaptive Synchronization Strategies for Automatic Accompaniment

José Echeveste developed several synchronization strategies in the framework of his PhD thesis. Their formalization is based on a dynamic real-time extension of the time map formalism, going beyond state-of-the-art where the largest body of literature on time maps is devoted to static functions, defined and known at all times before any manipulation is done. Only the latest work of Liang and Danneberg (2011) have considered dynamic time map in the synchronization problem. However their approach suffers from a consistency drawback: the convergence of the tempo depends on the events occurring during the catching trajectory. In our approach we have developed a lag-depend formulation of the catching trajectory, which is insensitive to the actual events. This adaptive strategy considers only the deviation in tempo and position and is otherwise context-independent, it ensures convergence both in position and tempo, and it is efficient: there is no need for a fine sampling clock to discretize the time evolution: as long as the prediction time map does not change, delays are computed only once using the accompaniment time map. Our approach is general enough to handle various important issues in automatic accompaniment: latency management, integration of non-constant tempo specifications in the score (*accelerando*, *ritardando*, *rubato*...), handling of missing events, etc. Synchronization strategies have been fully formalized in the PhD report of José Echeveste [8] together with a complete Antescofo core including other dynamic constructions.

7.6. Temporal objects for the design of reusable library in Antescofo

Composers develop their own idiosyncratic compositional language through their pieces. In addition, composers and sound engineers have to face drastically different performance set-up for the same piece. This situation advocates the development of new generic mechanisms to simplify the development of generic yet dedicated libraries in Antescofo. In cooperation with various composers (Marco Stroppa, Julia Blondeau, Jason Freeman, José Miguel Fernández, Yann Marez) we have introduced several new mechanisms in Antescofo to ease the building of dedicated yet reusable libraries of compositional pieces: extension of the functional language to include new control structures, introduction of *continuation combinators* making possible to start actions at the end of other durative actions, marshalling of Antescofo values, etc. The most notable ones are actor-based features to implement *temporal objects*. Object templates are specified and then instantiated at will. A temporal object encapsulates a local state; it can react to logical conditions; it offers instantaneous as well as durative methods; reaction to synchronous broadcast can be defined as well as exceptional condition handlers. These new features are currently tested in the development of new pieces and are expected to evolve following the feedbacks from these applications.

7.7. Embedding real-time audio computation in Antescofo

DSP processing in Antescofo is an experimental extension of the language started in 2014 and aimed at driving various DSP processing capabilities directly within Antescofo. DSP processors are defined directly in an Antescofo score, harnessing various signal processing libraries. These DSP processors are then dynamically

connected together using Antescofo audio links. Input and output channels are used to link these processors with the host environment while internal channels connect DSP among themselves. The connections are specified with a new kind of Antescofo actions, the patch. So, the connections can be changed dynamically in response to the events detected by the listening machine and can be synchronized using the expressive repertoire of synchronization strategies available in Antescofo. Ordinary Antescofo variables can be used to control the DSP computations, which add an additional level of dynamicity. Currently, FAUST and a few specific signal processors (notably FFT) can be defined. Several benefits result from this tight integration. The network of signal processors is heterogeneous, mixing DSP nodes specified with different tools. The network of signal processors can change dynamically in time following the result of a computation. This approach answers the shortcomings of fixed (static) dataflow models of the Max or PureData host environments. Signal processing is controlled at a symbolic level and can be guided, *e.g.* by information available in the augmented score (like position, expected tempo, etc.). The tight integration makes possible to specify, concisely and more effectively, finer and more precise control of the signal processing, at a lower computational cost. One example is the use of symbolic curve specification to specify variations of control parameters at sample rate. It makes it possible to embed sound analysis inside Antescofo as well. At last but not least, signal processing can be done more efficiently. For example, in the *remaking* of Boulez' piece *Antheme 2* there is an improvement of performance in time of 45 % compared to the original version with the audio effects managed in Max.

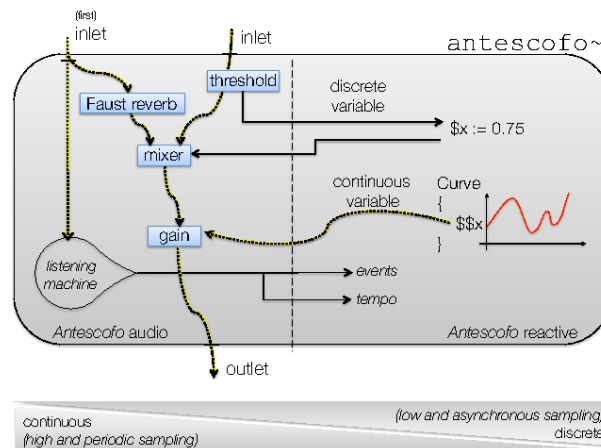


Figure 8. Articulation between DSP and the reactive engine.

The current work focuses on the development of a dedicated type system enabling a finer control of scheduling and audio buffer size, refining results previously developed in the cyclostatic scheduling of synchronous dataflow. Early results are published in [29].

7.8. Visualizing Timed and Hierarchical Code Structures

This work applies an information visualisation perspective to a set of revisions in the timeline-based representation of action items in *AscoGraph*, the dedicated user interface to Antescofo. Our contribution is twofold: (a) a design study of the proposed new model, and (b) a technical, algorithmic component. In the former, we show how our model relates to principles of information coherence and clarity, facility of seeking and navigation, hierarchical distinction and explicit linking. In the latter, we frame the problem of arranging action rectangles in a 2D space as a strip packing problem, with the additional constraint that the (horizontal) time coordinates of each block are fixed. We introduce three algorithms of increasing complexity for automatic arrangement, estimate their packing performance and analyse their strengths and weaknesses. We evaluate the systemic

improvements achieved and their applicability for other time-based datasets. Furthermore, algorithms for efficient automatic stacking of time-overlapping action blocks are developed, as well as mathematical proof for their time-coherency during dynamic visualizations.

Results are implemented in Section 6.2 and reported in [12] and [13].

7.9. Model-based Testing an Interactive Music System

We have been pursuing our studies on the application of model-based timed testing techniques to the interactive music system (IMS) Antescofo, in the context of the Phd of Clément Poncelet and in relation with the developments presented in Section 6.3 .

Several formal methods have been developed for automatic conformance testing of critical embedded software, with the execution of a real implementation under test (IUT, or black-box) in a testing framework, where carefully selected inputs are sent to the IUT and then the outputs are observed and analyzed. In conformance model-based testing (MBT), the input and corresponding expected outputs are generated according to formal models of the IUT and the environment. The case of IMS presents important originalities compared to other applications of MBT to realtime systems. On the one hand, the time model of IMS comprises several time units, including the wall clock time, measured in seconds, and the time of music scores, measured in number of beats relatively to a tempo. This situation raises several new problems for the generation of test suites and their execution. On the other hand, we can reasonably assume that a given mixed score of Antescofo specifies completely the expected timed behavior of the IMS, and compile automatically the given score into a formal model of the IUT's expected behavior, using an intermediate representation. This give a fully automatic test method, which is in contrast with other approaches which generally require experts to write the specification manually.

We have developed online and offline approaches to MBT for Antescofo. The offline approach relies on tools of the Uppaal suite [38], [37], using a translation of our models into timed automata. These results have been presented during the 30th ACM/SIGAPP Symposium On Applied Computing, track Software Verification and Testing [21] and an article describing this approach has been accepted for publication in the Journal of New Music Research. The online approach is based on a new virtual machine executing the models of score in intermediate representation (see Section 6.3).

7.10. Representation of Rhythm and Quantization

Rhythmic data are commonly represented by tree structures (rhythms trees) in assisted music composition environments, such as OpenMusic, due to the theoretical proximity of such structures with traditional musical notation. We are studying the application in this context of techniques and tools for processing tree structure, which were originally developed for other areas such as natural language processing, automatic deduction, Web data processing... We are particularly interested in two well established formalisms with solid theoretical foundations: tree automata and term rewriting.

Our first main contribution in that context is the development of a new framework for rhythm transcription, the problem of the generation, from a sequence of timestamped notes, *e.g.* a file in MIDI format, of a score in traditional music notation) – see Section 6.4 . This problem arises immediately as insoluble unequivocally: we shall calibrate the system to fit the musical context, balancing constraints of precision, or of simplicity / readability of the generated scores. We are developing in collaboration with Jean Bresson (Ircam) and Slawek Staworko (LINKS, currently on leave at University of Edinburgh) an approach based on algorithms for the enumeration of large sets of weighted trees (tree series), representing possible solutions to a problem of transcription. The implementation work is performed by Adrien Ycart, under a research engineer contract with Ircam. This work has been presented in [23].

Our second contribution, in collaboration with Prof. Masahiko Sakai (Nagoya University), is a proposal of a structural theory (equational system on rhythm trees) defining equivalence on rhythm notations [14], [16]. This approach can be used for example to generate, by transformation, different notations possible the same rate, with the ability to select in accordance with certain constraints. We have also conducted related work on the theory of term rewriting [17].

PARKAS Project-Team

6. New Results

6.1. Reasoning about C11 Program Transformations

Participants: Francesco Zappa Nardelli, Robin Morisset.

We have shown that the weak memory model introduced by the 2011 C and C++ standards does not permit many of common source-to-source program transformations (such as expression linearisation and "roach motel" reordering) that modern compilers perform and that are deemed to be correct. As such it cannot be used to define the semantics of intermediate languages of compilers, as, for instance, LLVM aimed to. We consider a number of possible local fixes, some strengthening and some weakening the model. We have evaluated the proposed fixes by determining which program transformations are valid with respect to each of the patched models. We have provided formal Coq proofs of their correctness or counterexamples as appropriate.

A paper on this work has been accepted in [18]. In collaboration with Viktor Vafeiadis (MPI-SWS, Germany) and Thibaut Balabonski (U. Paris Sud).

6.2. Language design on top of JavaScript

Participant: Francesco Zappa Nardelli.

This research project aims at improving the design of the JavaScript language. We propose a typed extension of JavaScript combining dynamic types, concrete types and like types to let developers pick the level of guarantee that is appropriate for their code. We have implemented our type system in the V8 JavaScript engine and we have explored the performance and software engineering benefits.

A paper on this work has been accepted in ECOOP 2015 [21].

With Gregor Richards (Waterloo University) and Jan Vitek (Northeastern University).

6.3. Synchronous Functional Language with Integer Clocks

Participant: Adrien Guatto.

Adrien Guatto defended his PhD thesis on the modular description of space/time tradeoffs at the language level. His thesis work extends the n-synchronous framework proposed by Cohen, Mandel, Plateau, Pouzet and others. Clocks now feature arbitrary positive integers that model bursty communication between subprograms: "integer clocks". The activation conditions of Lustre are revisited in this new setting to become "local time scales" that allow subprograms to perform several steps atomically relative to their context. The thesis details the integration of these features in a clock type system for a higher-order functional language, giving full formal treatment of its metatheory and compilation to finite-state digital circuits.

6.4. Fidelity in Real-Time Programming

Participants: Guillaume Baudart, Timothy Bourke.

In this work we study embedded systems with a significant mix of discrete reactive behaviours and 'physical' timing constraints. The idea is to make the most of the advantages of synchronous languages for precisely specifying discrete behaviours but to adapt or extend them to treat real-time constraints more abstractly, that is, without an *a priori* definition of an eventual sampling interval.

This year we concluded our study of the Loosely Timed-Triggered Architectures (LTTA) by developing simplified models of the underlying implementations and protocols. This enabled us to improve the protocols, simplify the correctness and performance arguments, and compare them to systems built using modern clock synchronization algorithms. We developed our models in the Zélus programming language which enables (instances of) them to be compiled for simulation and contributes to our work on better exploiting synchronous languages for real-time specification and analysis. This work was presented at the EMSOFT conference and a journal article has been submitted.

This year we also concluded our study of the Quasi-synchronous Approach to modelling real-time distributed systems. We formalized the relation between the discrete abstraction proposed by Paul Caspi and the real-time architectures for which it is intended. This enabled us to precisely state a correctness requirement for the abstraction and to show that it is sound for systems of two nodes (a typical case explored in other publications) but not for general systems of three or more nodes. Our formalization clarifies the relation between the causality of traces of the real-time system and the causality introduced by the synchronous abstraction. This enables us to state and show necessary and sufficient restrictions on the communication topologies and timing characteristics of systems to ensure soundness. A paper explaining this result has been drafted and will be submitted early in 2016.

6.5. Verified compilation of Lustre

Participants: Timothy Bourke, Marc Pouzet.

Synchronous dataflow languages and their compilers are increasingly used to develop safety-critical applications, like fly-by-wire controllers in aircraft and monitoring software for power plants. A striking example is the SCADE Suite tool of ANSYS/Esterel Technologies which is DO-178B/C qualified for the aerospace and defense industries. This tool allows engineers to develop and validate systems at the level of abstract block diagrams that are automatically compiled into executable code.

Formal modelling and verification in an interactive theorem prover can potentially complement the industrial certification of such tools to give very precise definitions of language features and increased confidence in their correct compilation; ideally, right down to the binary code that actually executes.

This year we picked up on previous work in the PARKAS team to develop a verified compiler for a Lustre/SCADE-like synchronous language. We focused on the critical and until now unresolved compiler stage that transforms dataflow equations into imperative code. We developed, in Coq, a prototype compiler for the core language (without modular resets or tuples) and showed its correctness with respect to a dataflow semantics based on functions from natural numbers to present or absent values. This required the development of a novel intermediate model for relating delayed dataflow streams to imperative memories in such a way that a critical induction could be stated and proved. We further showed how to justify a post-transformation optimization that is essential for the efficiency of clock-directed code generation. We are preparing a paper describing these results. Work continues on both semantic questions (existence of a semantics for well-type and well-clocked programs, treatment of resets, etc.) and compilation issues (integration with the verified CompCert compiler).

In collaboration with Pierre-Évariste Dagand (CNRS) and Lionel Reig (Collège de France).

PARSIFAL Project-Team

7. New Results

7.1. The Checkers Proof Certifier

Participants: Tomer Libal, Giselle Reis, Hichem Chihani.

We presented a system description [29] of the Checkers proof certifier, which implements some of the theoretical ideas developed in the ProofCert project. This version of the system is capable of certifying a subset of the E-Prover superposition theorem prover. The system is mainly written in λ Prolog with a proof importing module written in Ocaml. The system is designed to allow modularity when designing the semantical translations of proof systems. For this capacity, the system supports, J. A. Robinson's resolution and the paramodulation technique of G. Robinson and L. Wos. On top of that, minimal support for some inference rules of the E-Prover was added.

7.2. Regular Patterns in Second-Order Unification

Participant: Tomer Libal.

We presented a paper [33] detailing a higher-order pre-unification procedure with improved termination over existing procedures. The classic higher-order unification procedure was presented by G. Huet in 1975 and is still used as the main unification procedure for higher-order automated theorem provers. This procedure does not terminate. In this project we have investigated the reasons for that and have shown that by choosing a specific (but complete) search strategy, an additional set of non-unifiable problems can be detected. As an example, we have shown that all unification problems generated by the Leo-III theorem prover when proving Cantor's theorem are decided by this procedure, in contrast to the classical unification procedure.

7.3. Static guarantees for message-passing computation

Participant: Stéphane Graham-Lengrand.

LCF [79] is a proof-search architecture, where search strategies are programmed via an API and successful proof-search runs are guaranteed correct, relying on the use of an abstract type theorem. We adapted the approach and defined principles for message-passing software architectures (where modules interact by exchanging messages), with the objective of guaranteeing message provenance and integrity. The principles rely on abstract types to *sign* messages at no run-time cost, and more generally rely on type-checking to provide static guarantees (i.e. at compile-time) that the messages produced by a trusted piece of code will not be altered or faked by an untrusted piece of code. We developed this primarily for safe theorem proving architectures, but the approach can be applied to other software architectures where modules with different levels of trust interact.

7.4. Proof-search with quantifiers and theories

Participant: Stéphane Graham-Lengrand.

We published our approach to proof-search on quantified problems in presence of one theory [22], where we identify the specifications required of the theory for the proof-search process to be sound and complete. Theories with unification procedures or quantifier elimination procedures satisfy our specifications, where *constraint streams* and *constraint projections* play a key role. Interestingly enough, Bjorner and Janota [52] independently achieved a similar result with *model projections*. Our theory-generic approach allows a clear formulation of what it could mean to combine several quantifier-handling theories, hopefully generalising what the Nelson-Oppen combination technique does in a quantifier-free context. We recently obtained two new results towards this:

- First, the cumbersome, stream-querying, and backtracking mechanisms that were required to implement [22] have been re-expressed in a more satisfying message-passing computational framework.
- Second, we re-expressed the standard quantifier-free combination techniques, mentioned above, as a concurrent message-passing interaction between different theory-specific procedures, and simplified their proofs of correctness. This led to the major redesign of Psyche, mentioned above.

7.5. Realizability semantics of abstract focusing, formalized

Participant: Stéphane Graham-Lengrand.

In [21] we presented a parametric system for abstract focusing, building on Zeilberger’s work [87], and parametrically capturing classical and intuitionistic focused systems. We presented its semantics, building on Munch-Maccagnoni’s work [80], in terms of *abstract realizability models* (which were independently identified by Krivine). The goal was to emphasize the similarities and differences between focusing and realizability, in the way they exploit the *polarities* of formulae. The system and its semantics led to a substantial formalisation in the proof assistant Coq.

7.6. The Meta-Theory of Bisimulation-Up-To

Participants: Kaustuv Chaudhuri, Matteo Cimini, Dale Miller.

The method of proof by bisimulation has proved to be a very successful technique for showing the equivalence of processes. Unfortunately, in process calculi with infinite transition systems, such as in calculi with a replication operator, finding a bisimulation requires exploring an infinite search space, which moreover often tends to have rather intricate and complex structure. One way to combat this complexity—i.e., reduce the size of candidate bisimulation sets—is to identify redundancies among their members and then to replace redundant classes by unique inhabitants. This yields families of *bisimulation-up-to* proof methods that are parametric over the redundancy relation. For instance, if we consider bisimilarity itself as the redundancy, then we obtain *bisimulation up to bisimilarity*; with this relation, the singleton set $\{(!a, !!a)\}$ is a candidate set for showing that the processes $!a$ and $!!a$ are bisimilar, for example, when the bisimulation set with redundancies is infinite.

Since *a priori* there is no restriction on such redundancy relations, a key theoretical question is when a bisimulation-up-to relation is *sound*, i.e., that it is contained in a bisimulation. In the literature there have been a number of techniques proposed for showing soundness, but they often require the use of complex reasoning about lattices of fixed points. In [19] (CPP’15) we show how to use the built-in coinduction facilities of the Abella theorem prover to produce comparatively lightweight proofs of the soundness of many common bisimulation-up-to techniques for CCS and the π -calculus. A key feature of our approach is that we can use the facilities already provided by the Abella system for reasoning about the binding constructs for the π -calculus.

7.7. Characterizing Independence in Type Theory

Participants: Kaustuv Chaudhuri, Yuting Wang.

In formal proof languages based on type theory, it is often the case that a theorem is proved for a certain kind of typing context, but needs to be used in a different context. For example, theorems about natural numbers may be proved in an empty typing context, since the type of natural numbers contains no higher-order features (i.e., natural numbers are closed), but we may need to use these properties of natural numbers when reasoning about λ -terms in De Bruijn notation, where the typing context is non-empty. In such a situation, it is useful to automatically transport the existing theorems to the new kinds of contexts, since we know that the theorem in question cannot depend on the properties of λ -terms. While this example is rather trivial, it becomes non-trivial when theorems are proved about higher-order data structures, which are commonly encountered when reasoning about syntax with binding constructs.

One way to achieve such reuse automatically is a technique called *subordination*, which is based on analyzing the constructors for a certain type and defining syntactic criteria under which certain normal terms of one type can have subterms of another type. Unfortunately, the classical definition of subordination lacks a proof-theoretic justification, and has surprising properties in third-order (and higher) signatures.

In [36] (TLCA'15), we propose a proof-theoretic characterization of a kind of dual to subordination, called *independence*, that characterizes when normal terms of one type *cannot* contain subterms of another type. This is achieved by means of proving an inductive *strengthening* lemma about the signatures in the two-level logic approach. We also show how to automatically prove such lemmas in certain commonly encountered situations in the theorem prover Abella.

7.8. Disproving Non-Theorems with Saturating Search

Participants: Taus Brock-Nannestad, Kaustuv Chaudhuri.

High-performance automated reasoning techniques such as resolution and the inverse method are well suited for proving *true* conjectures, but are ill-behaved for *false* conjectures. For example, for a simple theory of even numbers that states that 0 is even and that $n + 2$ is even whenever n is even, it is obviously the case that the conjecture “3 is even” is unprovable, but the algorithm would loop forever proving “0 is even”, “2 is even”, “4 is even”, etc. This behavior is observed even in the best saturation-based (i.e., forward-reasoning) theorem provers.

In [25] (TABLEAUX'15), we show how to finitely constrain the search space of saturation-based theorem provers by the use of *unsound* extensions of the goal query. These unsound extensions, when combined with forward subsumption, guarantee that only a finite number of consequences would ever be constructed based on any goal query, so the proof search procedure is guaranteed to terminate. If a proof is found among them that does not use the unsound extensions, then we can succeed with that proof. If no proof is found, then we can soundly assert that the original goal query was also unprovable, since even a weakened version of it was unprovable. The only other possibility is that a proof is found using the unsound extension; in this case, we use the particular instance of unsoundness to refine the original unsound goal to prevent it from being found again, while maintaining the invariant that the search space is finite, and rerun the search. Since first-order logic is undecidable, we may need to repeat the refinement procedure indefinitely, but for many kinds of domains, particularly those arising from typed signatures (such as the even numbers example above), we do eventually find a saturating approximation that guarantees that the conjecture has no proof.

This algorithm has been implemented as part of the Mætning theorem prover explained in the section on Software above. We plan to extend it in the future with various automatic refinement heuristics.

7.9. Encoding Bigraph Structure with Subexponentials

Participants: Kaustuv Chaudhuri, Giselle Reis.

Bigraphs were proposed by Robin Milner as a model of ubiquitous computing, which is computation that is aware both of *location* and of *connections*. As a formalism it subsumes many other process calculi such as CCS and the π -calculus. However, it has a number of problems *qua* syntax because it is based on graphs and a complicated theory of composition. The biggest of these problems is how to implement it in a formal reasoning system.

In recent years, many members (and ex-members) of the Parsifal team have been experimenting with a variant of linear logic that has not just a single pair but an arbitrary family of exponential connectives that are arranged in a pre-order. Each such pair of *subexponentials* may admit or reject the structural properties of weakening and contraction. One benefit of subexponentials is that it allows for *querying* the *absence* of certain kinds of exponential formulas without requiring all non-exponential formulas to be deleted as a consequence, which is the issue with ordinary linear logic.

In [28] (LPAR'15), we show how to represent the structure of bigraphs in terms of a simple theory of linear logic with subexponentials (SEL). We show that our representation is adequate, i.e., that it respects the composition and juxtaposition operations on bigraphs. Moreover, we show how one can ask queries about the nesting of places in the representation without modifying it, which gives us a technical means of encoding bigraph reactions as well. Some of the details for bigraph reactions remain to be worked out in future work.

7.10. Encoding Additive Connectives with Multiplicatives and Subexponentials

Participant: Kaustuv Chaudhuri.

In a recent workshop on Linearity [55], we have published the formal proof (that was obtained in 2009) that linear logic with three subexponentials in a certain lattice is undecidable. An extended version of this paper was submitted to a special issue on Linearity in Mathematical Structures in Computer Science and was accepted in November 2015.

The preprint of that extended paper [41] gives a direct embedding of propositional MALL (multiplicative and additive linear logic) using only multiplicative connectives and five subexponentials. This means that the additive connectives are, in fact, redundant when we have multiplicatives and subexponentials. Moreover, in the first-order case this encoding is polynomial and focally adequate, which means that MALL can be simulated at the highest fidelity – at the level of individual inference rules.

7.11. Computation in Focused Intuitionistic Logic

Participants: Taus Brock-Nannestad, Nicolas Guenot, Daniel Gustafsson.

Focusing is a proof-theoretical technique for eliminating unnecessary nondeterminism in proofs. Because it cuts down on nondeterminism, focusing is particularly useful for directing proof search. Focusing thus plays a key role in explaining the meaning and behaviour of logic programs.

Despite this success in clarifying the operational semantics of logic programming, focusing has not been as widely studied in the Curry-Howard style “proofs as programs” interpretation. Early results in this area established that λ -calculi associated with the focused calculi LJT and LJQ had evaluation strategies corresponding to call-by-name and call-by-value respectively. For the LJF calculus — which contains both LJT and LJQ as fragments — no such correspondence was known.

In [27] (PPDP'15) we show how a proof-term assignment to (a variant of) Liang and Miller’s focused sequent calculus LJF permits a uniform treatment of the call-by-value and call-by-name reduction strategies of the λ -calculus, as well as combinations of these strategies. Additionally, we show how to extract an abstract machine from LJF by considering machine states as certain configurations of instances of the cut rule. The aforementioned correspondence extends to this setting, and we show that well-known abstract machines for call-by-value and call-by-name are in fact exactly the abstract machines that one gets when considering certain fragments of LJF.

In the seminal work of Paul Blain Levy, the call-by-push-value language was introduced as a way of subsuming the call-by-value and call-by-name strategies of the λ -calculus. It was later conjectured that call-by-push-value was simply implementing a notion of focusing, and indeed this turns out to be the case, as we show in the aforementioned paper.

7.12. Focused Linear Logic and the λ -calculus

Participants: Taus Brock-Nannestad, Nicolas Guenot.

Linear Logic enjoys strong symmetries inherited from classical logic while providing a constructive framework comparable to intuitionistic logic. However, the computational interpretation of sequent calculus presentations of linear logic remains problematic, mostly because of the many rule permutations allowed in the sequent calculus.

In focused variants of Linear Logic, most of these rule permutations are eliminated by the focusing restriction — during focusing, a single formula is decomposed eagerly, and the focus is passed down to its subformulas. Conversely, during inversion, all invertible connectives are decomposed. Moreover, this decomposition is made fully deterministic by keeping the connectives in question in a list, and only decomposing the first connective of this list.

The end result of this is that a focused proof in Linear Logic almost always has one particular formula singled out as the one that will be decomposed. Thus, somewhat curiously, focused Linear Logic behaves much more like an intuitionistic sequent calculus (where at all times there is a single “special” formula on the right hand side of the sequent) than a classical calculus.

In [26] (MFPS’15), we study a term assignment for a focused version of Multiplicative Exponential Linear Logic (MELL), and show how the focusing technique gives rise to a calculus that straightforwardly embeds both a linear variant of the λ -calculus, and a sequent-based formulation of Parigot’s $\lambda\mu$ -calculus.

7.13. There is no complete linear term rewriting system for propositional logic

Participant: Lutz Straßburger.

Recently, we observed that the set of all sound linear inference rules in propositional logic is already coNP-complete [84]. This means that every boolean tautology can be written as a (left-and right-) linear rewrite rule. This raises the question of whether there is a rewriting system on linear terms of propositional logic that is sound and complete for the set of all such rewrite rules. We have shown (in a joint work with Anupam Das) that, as long as reduction steps are polynomial-time decidable, such a rewriting system does not exist unless coNP=NP. This is published in [20].

7.14. A (Bi)linear Implementation of Strong Call-by-Value

Participant: Beniamino Accattoli.

The elegant theory of the call-by-value λ -calculus relies on closed terms and weak evaluation (i.e., not under abstractions) and it is well-known that the number of call-by-value β -steps is a reasonable cost model. When turning to open terms or strong evaluation—that are used for instance in the implementation of Coq—the operational theory breaks, and the call-by-value λ -calculus has to be extended with some additional rewriting rules. In a joint work with Sacerdoti Coen [18], a proposal for open/strong call-by-value, called *fireball calculus*, is studied from the point of view of cost models and abstract machines. First, it is shown that open terms introduce a new malicious behavior, making the study of cost models non-trivial. Second, it is shown that the number of β -steps in the fireball calculus is a reasonable cost model. Third, a new abstract machine is introduced and its overhead is shown to be linear with respect to the number of β -steps and the size of the initial term, providing a surprisingly efficient implementation scheme.

7.15. Implementations of Strong Call-by-Name, Revisited

Participant: Beniamino Accattoli.

The literature about abstract machines for the strong evaluation (i.e., possibly under abstraction) of the ordinary (i.e., call-by-name) λ -calculus is scarce. Essentially, there is a single, old work: Crégut’s abstract machine [60] (1990), that is an extension of Krivine abstract machine to compute full normal forms. Crégut studies the correctness of the machine by means of an explicit substitutions calculus. In this joint work with Barenbaum and Mazza [17], Crégut’s work is revisited and simplified in the extreme. An alternative, simpler machine is introduced, the *Strong Milner abstract machine*. Its correctness is studied via *linear substitution calculus*, a new approach to explicit substitutions developed by Accattoli and Kesner that is much simpler than Crégut’s approach. Moreover, a complexity analysis of the machine is provided: its overhead is shown to be linear in the number of steps in the linear substitution calculus and in the size of the initial term.

7.16. Foundational Proof Certificates

We have continued to explore a number of new aspects of framework we call *Foundational Proof Certificates* (FPCs). Besides having defined and implemented prototype checkers for FPCs in classical and intuitionistic logic [37] we have also extended the proof theory underlying numerous modal logics so that FPCs can be applied to modal logics [35]. We have also extended the notion of FPC to work also in the model checking setting [31]. In both the modal logic and model checking domains, the key to getting FPCs to work is to have descriptions of *focused proof systems* available for those logics.

Given that FPCs are declarative and semantically simple structures, it has been possible to find numerous applications of them outside the problem of simply checking them. It was shown, for example, that FPCs can be used to help define the semantics of the output from traditional theorem provers [23]. We have also used FPCs as proof outlines in order to define high-level tactics to direct proof search [24].

7.17. Multi-level Delimited Control

There has been a great deal of interest in recent years to providing interesting functional programming primitives that are based on classical logic and not just intuitionistic logic. Unfortunately, the standard sequent calculus proof theory for classical logic is far too chaotic to provide such a foundation. We have recently proposed adding to classical (linear) logic an assortment of *subexponentials* and to provide a rigid structure for their placement within formulas. This new framework allows for sequent calculus proof theory to provide to the functional programming paradigm the feature often called *multi-level delimited control* [32]. The main result in that paper is also noteworthy in that it shows how to build certain complex synthetic connectives even though the standard approach (using focusing proof systems) cannot be used.

PI.R2 Project-Team

6. New Results

6.1. Effects in proof theory and programming

Participants: Guillaume Claret, Pierre-Louis Curien, Hugo Herbelin, Étienne Miquey, Ludovic Patey, Pierre-Marie Pédrot, Yann Régis-Gianas, Alexis Saurin.

6.1.1. Axiom of dependent choice in classical arithmetic

In 2012, Hugo Herbelin showed that classical arithmetic in finite types extended with strong elimination of existential quantification proves the axiom of dependent choice. To get classical logic and choice together without being inconsistent is made possible first by constraining strong elimination of existential quantification to proofs that are essentially intuitionistic and secondly by turning countable universal quantification into an infinite conjunction of classical proofs evaluated along a call-by-need evaluation strategy so as to extract from them intuitionistic contents that complies to the intuitionistic constraint put on strong elimination of existential quantification. Étienne Miquey has been working on a sequent-calculus version of this system, using Danvy's methodology of semantic artifacts, to progressively reduce the consistency of such a system to the normalisation of Girard-Reynold's system F. To achieve this goal, he incidentally proposed a way to get a dependently-typed sequent calculus, as well as a method to type a state-and-continuation-passing style translation of call-by-need calculus.

6.1.2. The computational contents of completeness proofs

Hugo Herbelin worked on the computational content of Gödel's completeness theorem, developing a proof with side-effects suitable for normalisation-by-evaluation.

6.1.3. Gödel's functional interpretation

Pierre-Marie Pédrot extended the proof-as-program interpretation of Gödel's Dialectica translation to the fully dependent setting, including dependent elimination [17].

6.1.4. Logical foundations of call-by-need evaluation

Alexis Saurin and Pierre-Marie Pédrot extended their reconstruction of call-by-need based on linear head reduction with control. They showed how linear head reduction could be adapted to the $\lambda\mu$ -calculus. This classical linear head reduction lifts the usual properties of the intuitionistic one (with respect to σ -equivalence) to the $\lambda\mu$ -calculus (and its σ -equivalence already formulated by Olivier Laurent in his PhD thesis). Moreover, they showed that substitution sequences of the $\lambda\mu$ -calculus linear head reduction are in correspondence with the classical Krivine abstract machine substitution sequences, validating the known fact that the KAM implements linear head reduction. In a second step, they could lift to the $\lambda\mu$ -calculus their three-step transformation from linear head reduction to call-by-need, and study the correspondence with Ariola, Herbelin and Saurin's classical call-by-need. This work appeared as one of the chapters of Pierre-Marie Pédrot's thesis and has been accepted for publication at ESOP'16 [30].

6.1.5. Call-by-name forcing

Pierre-Marie Pédrot studied variants of the forcing construction by decomposing it through call-by-push-value. In particular, the by-name decomposition behaves much more nicely w.r.t. the computational content of proofs and is a candidate for a dependently-typed extension. This work is partially reported on in his PhD [17].

6.1.6. A theory of effects and resources

In joint work with Marcelo Fiore and Guillaume Munch-Maccagnoni, Pierre-Louis Curien considered the Curry-Howard-Lambek correspondence for effectful computation and resource management, specifically proposing polarised calculi together with presheaf-enriched adjunction models as the starting point for a comprehensive semantic theory relating logical systems, typed calculi, and categorical models in this context. Our thesis is that the combination of effects and resources should be considered orthogonally. Model theoretically, this leads to an understanding of our categorical models from two complementary perspectives: (i) as a linearisation of CBPV (Call-by-Push-Value) adjunction models, and (ii) as an extension of linear/non-linear adjunction models with an adjoint resolution of computational effects. When the linear structure is cartesian and the resource structure is trivial, we recover Levy’s notion of CBPV adjunction model, while when the effect structure is trivial, we have Benton’s linear/non-linear adjunction models. Further instances of our model theory include the dialogue categories with a resource modality of Melliès and Tabareau, and the Enriched Effect Calculus models of Egger, Møgelberg and Simpson. Our development substantiates the approach by providing a lifting theorem of linear models into cartesian ones. To each of our categorical models we systematically associate a typed term calculus, each of which corresponds to a variant of the sequent calculi LJ (Intuitionistic Logic) or ILL (Intuitionistic Linear Logic). The adjoint resolution of effects corresponds to polarisation whereby, syntactically, types locally determine a strict or lazy evaluation order and, semantically, the associativity of cuts is relaxed. In particular, our results show that polarisation provides a computational interpretation of CBPV in direct style. Further, we characterise depolarised models: those where the cut is associative, and where the evaluation order is unimportant. This work will be presented at POPL 2016 [26].

6.1.7. Coq as a programming language with effects

As part of his PhD thesis, Guillaume Claret defined a notion of effectful interactive computation as an embedded DSL in Coq (in the spirit of the works on algebraic effects), and used it to implement a web server. It is equipped with a dual notion of effectful interactive execution context. Using these two notions together, Guillaume Claret is able to specify and reason about interactive programs inside Coq. He submitted several papers about this line of work: one has been published [32], others will be part of his PhD manuscript.

6.2. Reasoning and programming with infinite data

Participants: Amina Doumane, Alexis Saurin, Pierre-Marie Pédro, Yann Régis-Gianas.

This theme is part of the ANR project Rapido (see the National Initiatives section).

6.2.1. Interactive semantics for logic fixed-points and infinitary logics.

Amina Doumane and Alexis Saurin, in a joint work with David Baelde published at CSL 2015 [24], developed a game-semantics of $\mu MALL$ (Multiplicative Additive Linear Logic with least and greatest fixpoints).

This interactive semantics was worked out in computational ludics, benefitting from both the work by Clairambault on a HO style game semantics for an intuitionistic logic with least and greatest fixpoints and from the flexibility of Terui’s computational ludics (in particular its ability to consider designs with cuts).

This framework is built around the notion of design, which can be seen as an analogue of the strategies of game semantics. The infinitary nature of designs makes them particularly well suited for representing computations over infinite data. We provided $\mu MALL$ with a denotational semantics (that is invariant by cut-elimination), interpreting proofs by designs and formulas by particular sets of designs called behaviours. Then a completeness result for a specific class of designs is proved, the class of “essentially finite designs”, which are those designs performing a finite computation followed by a copycat. On the way to the previous completeness result, we investigate semantic inclusion, proving its decidability (given two formulas A and B , one can decide whether the semantics of A is included in the semantics of B) and completeness (if semantic inclusion holds, the corresponding implication is provable in $\mu MALL$).

6.2.2. Proof theory of circular proofs

In a collaboration with David Baelde, Amina Doumane and Alexis Saurin developed further the theory of infinite proofs. Studying the proof theory of circular proofs on MALL, they established a result of focalisation for these infinite proofs. The usual result of focalisation for linear logic can actually be extended to circular proofs but, contrarily to μ MALL where fixed-points operators can be given an arbitrary polarity, the least fixed-point must be set to be a positive construction and the greatest fixed-points to be negative, which is consistent with intuition from programming with inductive and co-inductive datatypes. An interesting phenomenon arising with focalisation is that some infinite but regular proofs may not have any regular focused proofs. This is similar to what happens for cut-elimination of regular proofs.

Works on cut-elimination for circular proofs are still ongoing.

6.2.2.1. Automata theory meets proof theory: proof certificates for Büchi inclusion

In a joint work with David Baelde and Lucca Hirschi, Amina Doumane and Alexis Saurin carried out a proof-theoretical investigation of the linear-time μ -calculus, proposing well-structured proof systems and showing constructively that they are complete for inclusions of Büchi automata suitably encoded as formulas.

They do so in a way that combines the advantages of two lines of previous work: Kaivola gave a proof of completeness for an axiomatisation that amounts to a finitary proof system, but his proof is non-constructive and yields no reasonable procedure. On the other hand, Dax, Hofmann and Lange recently gave a deductive system that is appropriate for algorithmic proof search, but their proofs require a global validity condition and do not have a well understood proof theory.

They work with well-structured proof systems, effectively constructing proofs in a finitary sequent calculus that enjoys local correctness and cut elimination.

This involves an intermediate circular proof system in which one can obtain proofs for all inclusions of parity automata, by adapting Safra's construction. In order to finally obtain finite proofs of Büchi inclusions, a translation result from circular to finite proofs is designed.

6.3. Effective higher dimensional algebra

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos, Jovana Obradović.

6.3.1. Rewriting methods for Artin monoids

With Stéphane Gaussent (ICJ, Univ. Saint-Étienne), Yves Guiraud and Philippe Malbos have used higher-dimensional rewriting methods for the study of Artin monoids, a class of monoids that is fundamental in algebra and geometry. This work formulates in a common language several known results in combinatorial group theory: one by Tits about the fundamental group of a graph associated to an Artin monoid [76], and one by Deligne about the actions of Artin monoids on categories [58], both originally proved by geometrical and non-constructive methods. An improved completion procedure, called the homotopical completion-reduction procedure (see also [8]), is formalised and used to give constructive proofs of (improved versions of) both theorems. This work has been published in *Compositio Mathematica* [19] and has been implemented in a Python library (<http://www.pps.univ-paris-diderot.fr/~guiraud/cox/cox.zip>).

6.3.2. Rewriting and Garside theory

Yves Guiraud has collaborated with Patrick Dehornoy (LNO, Univ. Caen) to develop an axiomatic setting for monoids with a special notion of quadratic normalisation map with good computational properties. This theory generalises the normalisation procedure known for monoids that admit a special family of generators called a Garside family [57] to a much wider class that also includes the plactic monoids. It is proved that good quadratic normalisation maps correspond to quadratic convergent presentations, together with a sufficient condition for this to happen, based on the shape of the normalisation paths on length-three words. This work has been submitted for publication to the *Journal de l'École Polytechnique — Mathématiques* [44].

Building on this last article, Yves Guiraud currently collaborates with Matthieu Picantin (Automates team, LIAFA, Univ. Paris 7) to generalise the main results of [19] to monoids with a Garside family. This will allow an extension of the field of application of the rewriting methods to other geometrically interesting classes of monoids, such as the dual braid monoids.

6.3.3. Higher-dimensional linear rewriting

With Eric Hoffbeck (LAGA, Univ. Paris 13), Yves Guiraud and Philippe Malbos have introduced in [64] the setting of linear polygraphs to formalise a theory of linear rewriting, generalising Gröbner bases. They have adapted the computational method of [7] to compute polygraphic resolutions of associative algebras, with applications to the decision of the Koszul homological property. They are currently engaged into a major overhaul of this work, whose main goal is to ease the adaptation of the results to other algebraic varieties, like commutative algebras or Lie algebras.

6.3.4. Theory of reduction operators

Cyrille Chenavier, supervised by Yves Guiraud and Philippe Malbos, explores the use of Berger's theory of reduction operators [50] to design new rewriting methods in algebra. In [42], he proposed a construction of a contracting homotopy for the Koszul complex of an algebra (a complex characterising the homological property of Koszulness): when an algebra admits a side-confluent presentation (a strong hypothesis of confluence), he gave a candidate for the contracting homotopy, built using specific representations of confluence algebras; when the presentation satisfies an additional condition, called the extra-condition, it turns out that this candidate works.

6.3.5. Rewriting methods for coherence

In [45], Maxime Lucas, supervised by Yves Guiraud and Pierre-Louis Curien, has applied the rewriting techniques of [65] to prove coherence theorems for bicategories and pseudofunctors. He obtained a coherence theorem for pseudonatural transformations thanks to a new theoretical result, improving on the former techniques, that relates the properties of rewriting in 1- and 2-categories.

6.3.6. Wiring structure of operads and operad-like structures

Building on recent ideas of Marcelo Fiore on the one hand, and of François Lamarche on the other hand, Pierre-Louis Curien and Jovana Obradović developed a syntactic approach, using some of the kit of Curien-Herbelin's duality of computation and its polarised versions by Munch and Curien, to the definition of various structures that have appeared in algebra under the names of operads, cyclic operads, dioperads, properads, modular and wheeled operads, permutads, etc. These structures are defined in the literature in different flavours. The goal is to formalise the proofs of equivalence between these different styles of definition. This work is completed for cyclic operads and was presented at the conference Category Theory 2015 in Aveiro [43]. Further work will be to make these proofs modular, so as not to repeat them for each variation of the notion of operad.

6.3.7. A graphical proof of the Bénabou-Roubaud theorem

As a substantial development of reasoning with string diagrams, Jovana Obradović gave a complete proof of the Bénabou-Roubaud monadic descent theorem in [47]. One of the essential points concerning Grothendieck's original approach to descent theory consists of identifying the class of effective descent morphisms for a given fibration. In the special case of a bifibration satisfying Beck-Chevalley condition, Bénabou and Roubaud have given such a characterisation by means of monadicity. Due to the technically complicated calculations involving Grothendieck's cocycle condition, the categorical equivalence which reflects the comparison of the descent in fibered categories with monadic descent is usually not worked out in complete detail in the literature. Jovana Obradović linked the monadic and the original viewpoint via another possible definition of the category of descent data. This intermediate step, due to Janelidze and Tholen, involves constructions in internal categories and it provides an example on how one can stay in the world of string diagrams even when dealing with morphisms which do not have an explicit string diagram definition.

6.4. Incrementality

Participants: Yann Régis-Gianas, Lourdes Del Carmen González Huesca, Thibaut Girka.

An optimisation to perform incremental computations was developed by Lourdes del Carmen González Huesca and Yann Régis-Gianas, providing a mechanism to achieve efficiency. Incrementality as a way to propagate an input change into a corresponding output change is guided by formal change descriptions over terms and dynamic differentiation of functions. The data-changes are represented by displaceable types, a general framework to displace terms directed by types. An extension of the simply-typed lambda-calculus with differentials and partial derivatives offers a language to reason about incrementality. The basic system, λ -diff, was enriched with expressions for fixed-points and data-types together with their corresponding derivatives to analyse incrementality over them. The above results are reported in the second part of Lourdes González Huesca PhD thesis [16].

In collaboration with Paolo Giarrusso and Yufei Cai (Univ Marburg, Allemagne), Yann Régis-Gianas developed a new method to incrementalise higher-order programs using formal derivatives and static caching. A paper is in preparation.

In collaboration with David Mentré (Mitsubishi), Thibaut Girka and Yann Régis-Gianas designed and certified a new algorithm for correlating program generation: such a program is used to characterise the differences between two close programs. (Therefore, a correlating program is a good input for an incremental static analyser.) Before their work, only one algorithm existed in the literature and it was unsound. The new algorithm is sound and certified in Coq. This work has been published in the ATVA conference. Thibaut Girka has presented this work [33] at ATVA 2015.

In collaboration with David Mentré (Mitsubishi), Thibaut Girka and Yann Régis-Gianas are developing a theoretical framework to define a notion of differential operational semantics: a general mathematical object to characterise the difference of behavior of two close programs.

6.5. Metatheory and development of Coq

Participants: Pierre-Louis Curien, Hugo Herbelin, Pierre Letouzey, Yann Régis-Gianas, Matthieu Sozeau.

6.5.1. Models of type theory

Simplicial sets and their extensions as Kan complexes can serve as models of homotopy type theory. Hugo Herbelin extended his concrete type-theoretic formalisation of semi-simplicial sets [20] to simplicial sets.

6.5.2. Unification

Matthieu Sozeau is working in collaboration with Beta Ziliani (PhD at MPI-Saarbrücken, now assistant professor at Cordoba, Argentina) on formalising the unification algorithm used in Coq, which is central for working with advanced type inference features like Canonical Structures. This is the first precise formalisation of all the rules of unification including the ones used for canonical structure resolution and universes. The presentation includes a careful study of the heuristics used in the existing Coq algorithms, which can be added or removed from the new implementation modularly. This work has been presented at the ICFP'15 conference [31].

6.5.3. Nominal techniques

Matthieu Sozeau cosupervised the internship of Gabriel Lewertowski with Nicolas Tabareau (Ascola team, Nantes), on the development of a library for nominal reasoning in Coq/Ssreflect. The goal of this internship was to study the use of nominal sets to ease the formalisation of programming language (meta-)theory. A library based on the Mathematical Components formalisation of finite sets and effective quotients was built, providing generic definitions of substitution and elimination operators for simple descriptions of programming language syntax as a grammar. This work was done in collaboration with Assia Mahboubi (Specfun) and Cyril Cohen (Marelle). It forms the basis for the formalisation of cubical type theory, a new type theory using name abstraction that implements an axiom-free version of Homotopy Type Theory.

POLSYS Project-Team

6. New Results

6.1. Fundamental algorithms and structured polynomial systems

6.1.1. On the complexity of the F_5 Gröbner basis algorithm

We study the complexity of Gröbner bases computation, in particular in the generic situation where the variables are in simultaneous Noether position with respect to the system.

We give a bound on the number of polynomials of degree d in a Gröbner basis computed by F_5 algorithm in this generic case for the grevlex ordering (which is also a bound on the number of polynomials for a reduced Gröbner basis, independently of the algorithm used). Next, we analyse more precisely the structure of the polynomials in the Gröbner bases with signatures that F_5 computes and use it to bound the complexity of the algorithm.

Our estimates show that the version of F_5 we analyse, which uses only standard Gaussian elimination techniques, outperforms row reduction of the Macaulay matrix with the best known algorithms for moderate degrees, and even for degrees up to the thousands if Strassen's multiplication is used. The degree being fixed, the factor of improvement grows exponentially with the number of variables.

6.1.2. On the complexity of computing Gröbner bases for weighted homogeneous systems

Solving polynomial systems arising from applications is frequently made easier by the structure of the systems. Weighted homogeneity (or quasi-homogeneity) is one example of such a structure: given a system of weights $W = (w_1, \dots, w_n)$, W -homogeneous polynomials are polynomials which are homogeneous w.r.t the weighted degree $\deg(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) = \sum_{i=1}^n w_i \alpha_i$. Gröbner bases for weighted homogeneous systems can be computed by adapting existing algorithms for homogeneous systems to the weighted homogeneous case. In [6], we show that in this case, the complexity estimate for Algorithm F_5 $\left(\binom{n+d_{\max}-1}{d_{\max}}\right)^\omega$ can be divided by a factor $(\prod_{i=1}^n w_i)^\omega$. For zero-dimensional systems, the complexity of Algorithm FGLM nD^ω (where D is the number of solutions of the system) can be divided by the same factor $(\prod_{i=1}^n w_i)^\omega$. Under genericity assumptions, for zero-dimensional weighted homogeneous systems of W -degree (d_1, \dots, d_n) , these complexity estimates are polynomial in the weighted Bézout bound $\prod_{i=1}^n d_i / \prod_{i=1}^n w_i$. Furthermore, the maximum degree reached in a run of Algorithm F_5 is bounded by the weighted Macaulay bound $\sum_{i=1}^n (d_i - w_i) + w_n$, and this bound is sharp if we can order the weights so that $w_n = 1$. For overdetermined semi-regular systems, estimates from the homogeneous case can be adapted to the weighted case. We provide some experimental results based on systems arising from a cryptography problem and from polynomial inversion problems. They show that taking advantage of the weighted homogeneous structure yields substantial speed-ups, and allows us to solve systems which were otherwise out of reach.

6.1.3. Linear Algebra for Computing Gröbner Bases of Linear Recursive Multidimensional Sequences

Sakata generalized the Berlekamp – Massey algorithm to n dimensions in 1988. The Berlekamp – Massey – Sakata (BMS) algorithm can be used for finding a Gröbner basis of a 0-dimensional ideal of relations verified by a table. We investigate this problem using linear algebra techniques, with motivations such as accelerating change of basis algorithms (FGLM) or improving their complexity. In [12], we first define and characterize multidimensional linear recursive sequences for 0-dimensional ideals. Under genericity assumptions, we propose a randomized preprocessing of the table that corresponds to performing a linear change of coordinates on the polynomials associated with the linear recurrences. This technique then essentially reduces our problem to using the efficient 1-dimensional Berlekamp – Massey (BM) algorithm. However, the number of probes to the table in this scheme may be elevated. We thus consider the table in

the *black-box* model: we assume probing the table is expensive and we minimize the number of probes to the table in our complexity model. We produce an FGLM-like algorithm for finding the relations in the table, which lets us use linear algebra techniques. Under some additional assumptions, we make this algorithm adaptive and reduce further the number of table probes. This number can be estimated by counting the number of distinct elements in a multi-Hankel matrix (a multivariate generalization of Hankel matrices); we can relate this quantity with the *geometry* of the final staircase. Hence, in favorable cases such as convex ones, the complexity is essentially linear in the size of the output. Finally, when using the LEX ordering, we can make use of fast structured linear algebra similarly to the Hankel interpretation of Berlekamp – Massey.

6.1.4. Nearly optimal computations with structured matrices

In [9] we estimate the Boolean complexity of multiplication of structured matrices by a vector and the solution of nonsingular linear systems of equations with these matrices. We study four basic and most popular classes, that is, Toeplitz, Hankel, Cauchy and Vandermonde matrices, for which the cited computational problems are equivalent to the task of polynomial multiplication and division and polynomial and rational multipoint evaluation and interpolation. The Boolean cost estimates for the latter problems have been obtained by Kirrinnis in [10], except for rational interpolation. We supply them now as well as the Boolean complexity estimates for the important problems of multiplication of transposed Vandermonde matrix and its inverse by a vector. All known Boolean cost estimates for such problems rely on using Kronecker product. This implies the d -fold precision increase for the d -th degree output, but we avoid such an increase by relying on distinct techniques based on employing FFT. Furthermore we simplify the analysis and make it more transparent by combining the representations of our tasks and algorithms both via structured matrices and via polynomials and rational functions. This also enables further extensions of our estimates to cover Trummer’s important problem and computations with the popular classes of structured matrices that generalize the four cited basic matrix classes, as well as the transposed Vandermonde matrices. It is known that the solution of Toeplitz, Hankel, Cauchy, Vandermonde, and transposed Vandermonde linear systems of equations is generally prone to numerical stability problems, and numerical problems arise even for multiplication of Cauchy, Vandermonde, and transposed Vandermonde matrices by a vector. Thus our FFT-based results on the Boolean complexity of these important computations could be quite interesting because our estimates are reasonable even for more general classes of structured matrices, showing rather moderate growth of the complexity as the input size increases.

6.2. Solving Polynomial Systems over the Reals and Applications

6.2.1. Probabilistic Algorithm for Computing the Dimension of Real Algebraic Sets

Let $f \in \mathbb{Q}[X_1, \dots, X_n]$ be a polynomial of degree D . We consider the problem of computing the real dimension of the real algebraic set defined by $f = 0$. Such a problem can be reduced to quantifier elimination. Hence it can be tackled with Cylindrical Algebraic Decomposition within a complexity that is doubly exponential in the number of variables. More recently, denoting by d the dimension of the real algebraic set under study, deterministic algorithms running in time $D^{O(d(n-d))}$ have been proposed. However, no implementation reflecting this complexity gain has been obtained and the constant in the exponent remains unspecified. In [11], we design a probabilistic algorithm which runs in time which is essentially cubic in $D^{d(n-d)}$. Our algorithm takes advantage of genericity properties of polar varieties to avoid computationally difficult steps of quantifier elimination. We also report on a first implementation. It tackles examples that are out of reach of the state-of-the-art and its practical behavior reflects the complexity gain.

6.2.2. Real root finding for determinants of linear matrices

Let A_0, A_1, \dots, A_n be given square matrices of size m with rational coefficients. The paper [7] focuses on the exact computation of one point in each connected component of the real determinantal variety $\{x \in \mathbb{R}^n : \det(A_0 + x_1 A_1 + \dots + x_n A_n) = 0\}$. Such a problem finds applications in many areas such as control theory, computational geometry, optimization, etc. Using standard complexity results this problem can be solved using $m^{O(n)}$ arithmetic operations. Under some genericity assumptions on the coefficients of the

matrices, we provide in an algorithm solving this problem whose runtime is essentially quadratic in $\binom{n+m}{n}^3$. We also report on experiments with a computer implementation of this algorithm. Its practical performance illustrates the complexity estimates. In particular, we emphasize that for subfamilies of this problem where m is fixed, the complexity is polynomial in n .

6.2.3. Real root finding for rank defects in linear Hankel matrices

Let H_0, \dots, H_n be $m \times m$ matrices with entries in \mathbb{Q} and Hankel structure, i.e. constant skew diagonals. We consider the linear Hankel matrix $H(X) = H_0 + X_1 H_1 + \dots + X_n H_n$ and the problem of computing sample points in each connected component of the real algebraic set defined by the rank constraint $\text{rank}(H(X)) \leq r$, for a given integer $r \leq m - 1$. Computing sample points in real algebraic sets defined by rank defects in linear matrices is a general problem that finds applications in many areas such as control theory, computational geometry, optimization, etc. Moreover, Hankel matrices appear in many areas of engineering sciences. Also, since Hankel matrices are symmetric, any algorithmic development for this problem can be seen as a first step towards a dedicated exact algorithm for solving semi-definite programming problems, i.e. linear matrix inequalities. Under some genericity assumptions on the input (such as smoothness of an incidence variety), we design in [18] a probabilistic algorithm for tackling this problem. It is an adaptation of the so-called critical point method that takes advantage of the special structure of the problem. Its complexity reflects this: it is essentially quadratic in specific degree bounds on an incidence variety. We report on practical experiments and analyze how the algorithm takes advantage of this special structure. A first implementation outperforms existing implementations for computing sample points in general real algebraic sets: it tackles examples that are out of reach of the state-of-the-art.

6.2.4. Optimizing a Parametric Linear Function over a Non-compact Real Algebraic Variety

In [17], we consider the problem of optimizing a parametric linear function over a non-compact real trace of an algebraic set. Our goal is to compute a representing polynomial which defines a hypersurface containing the graph of the optimal value function. Rostalski and Sturmfels showed that when the algebraic set is irreducible and smooth with a compact real trace, then the least degree representing polynomial is given by the defining polynomial of the irreducible hypersurface dual to the projective closure of the algebraic set. First, we generalize this approach to non-compact situations. We prove that the graph of the opposite of the optimal value function is still contained in the affine cone over a dual variety similar to the one considered in compact case. In consequence, we present an algorithm for solving the considered parametric optimization problem for generic parameters' values. For some special parameters' values, the representing polynomials of the dual variety can be identically zero, which give no information on the optimal value. We design a dedicated algorithm that identifies those regions of the parameters' space and computes for each of these regions a new polynomial defining the optimal value over the considered region.

6.2.5. Bounds for the Condition Number of Polynomials Systems with Integer Coefficients

Polynomial systems of equations are a central object of study in computer algebra. Among the many existing algorithms for solving polynomial systems, perhaps the most successful numerical ones are the homotopy algorithms. The number of operations that these algorithms perform depends on the condition number of the roots of the polynomial system. Roughly speaking the condition number expresses the sensitivity of the roots with respect to small perturbation of the input coefficients. A natural question to ask is how can we bound, in the worst case, the condition number when the input polynomials have integer coefficients? In [19] we address this problem and we provide effective bounds that depend on the number of variables, the degree and the maximum coefficient bitsize of the input polynomials. Such bounds allows to estimate the bit complexity of the algorithms that depend on the separation bound, like the homotopy algorithms, for solving polynomial systems.

6.2.6. Nearly Optimal Refinement of Real Roots of a Univariate Polynomial

In [10] we assume that a real square-free polynomial A has a degree d , a maximum coefficient bitsize τ and a real root lying in an isolating interval and having no nonreal roots nearby (we quantify this assumption). Then, we combine the *Double Exponential Sieve* algorithm (also called the *Bisection of the Exponents*), the

bisection, and Newton iteration to decrease the width of this inclusion interval by a factor of $t = 2^{-L}$. The algorithm has Boolean complexity $\tilde{O}_B(d^2\tau + dL)$. Our algorithms support the same complexity bound for the refinement of r roots, for any $r \leq d$.

6.2.7. Accelerated Approximation of the Complex Roots and Factors of a Univariate Polynomial

The known algorithms approximate the roots of a complex univariate polynomial in nearly optimal arithmetic and Boolean time. They are, however, quite involved and require a high precision of computing when the degree of the input polynomial is large, which causes numerical stability problems. We observe that these difficulties do not appear at the initial stages of the algorithms, and in [8] we extend one of these stages, analyze it, and avoid the cited problems, still achieving the solution within a nearly optimal complexity estimates, provided that some mild initial isolation of the roots of the input polynomial has been ensured. The resulting algorithms promise to be of some practical value for root-finding and can be extended to the problem of polynomial factorization, which is of interest on its own right. We conclude with outlining such an extension, which enables us to cover the cases of isolated multiple roots and root clusters.

6.2.8. Polynomial Interrupt Timed Automata

Interrupt Timed Automata (ITA) form a subclass of stopwatch automata where reachability and some variants of timed model checking are decidable even in presence of parameters. They are well suited to model and analyze real-time operating systems. Here we extend ITA with polynomial guards and updates, leading to the class of polynomial ITA (polITA). In [13], we prove that reachability is decidable in 2EXPTIME on polITA, using an adaptation of the cylindrical algebraic decomposition algorithm for the first-order theory of reals using symbolic computation. Compared to previous approaches, our procedure handles parameters and clocks in a unified way. We also obtain decidability for the model checking of a timed version of CTL and for reachability in several extensions of polITA.

6.3. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory

6.3.1. Polynomial-Time Algorithms for Quadratic Isomorphism of Polynomials: The Regular Case

Let $\mathbf{f} = (f_1, \dots, f_m)$ and $\mathbf{g} = (g_1, \dots, g_m)$ be two sets of $m \geq 1$ nonlinear polynomials in $\mathbb{K}[x_1, \dots, x_n]$ (\mathbb{K} being a field). In [3], we consider the computational problem of finding – if any – an invertible transformation on the variables mapping \mathbf{f} to \mathbf{g} . The corresponding equivalence problem is known as *Isomorphism of Polynomials with one Secret* (IP1S) and is a fundamental problem in multivariate cryptography. Amongst its applications, we can cite Graph Isomorphism (GI) which reduces to equivalence of cubic polynomials with respect to an invertible linear change of variables, according to Agrawal and Saxena. The main result is a randomized polynomial-time algorithm for solving IP1S for quadratic instances, a particular case of importance in cryptography. To this end, we show that IP1S for quadratic polynomials can be reduced to a variant of the classical module isomorphism problem in representation theory. We show that we can essentially *linearize* the problem by reducing quadratic-IP1S to test the orthogonal simultaneous similarity of symmetric matrices; this latter problem was shown by Chistov, Ivanyos and Karpinski (ISSAC 1997) to be equivalent to finding an invertible matrix in the linear space $\mathbb{K}^{n \times n}$ of $n \times n$ matrices over \mathbb{K} and to compute the square root in a certain representation in a matrix algebra. While computing square roots of matrices can be done efficiently using numerical methods, it seems difficult to control the bit complexity of such methods. However, we present exact and polynomial-time algorithms for computing a representation of the square root of a matrix in $\mathbb{K}^{n \times n}$, for various fields (including finite fields), as a product of two matrices. Each coefficient of these matrices lie in an extension field of \mathbb{K} of polynomial degree. We then consider #IP1S, the counting version of IP1S for quadratic instances. In particular, we provide a (complete) characterization of the automorphism group of homogeneous quadratic polynomials. Finally, we also consider the more general *Isomorphism of Polynomials* (IP) problem where we allow an invertible linear transformation on the variables *and* on the set

of polynomials. A randomized polynomial-time algorithm for solving IP when $\mathbf{f} = (x_1^d, \dots, x_n^d)$ is presented. From an algorithmic point of view, the problem boils down to factoring the determinant of a linear matrix (i.e. a matrix whose components are linear polynomials). This extends to IP a result of Kayal obtained for PolyProj.

6.3.2. Factoring $N = p^r q^s$ for Large r and s

Boneh *et al.* showed at Crypto 99 that moduli of the form $N = p^r q$ can be factored in polynomial time when $r \simeq \log p$. Their algorithm is based on Coppersmith's technique for finding small roots of polynomial equations. In [15] we show that $N = p^r q^s$ can also be factored in polynomial time when r or s is at least $(\log p)^3$; therefore we identify a new class of integers that can be efficiently factored. We also generalize our algorithm to moduli with k prime factors $N = \prod_{i=1}^k p_i^{r_i}$; we show that a non-trivial factor of N can be extracted in polynomial-time if one of the exponents r_i is large enough.

6.3.3. On the Complexity of the BKW Algorithm on LWE

This work [1] presents a study of the complexity of the Blum–Kalai–Wasserman (BKW) algorithm when applied to the Learning with Errors (LWE) problem, by providing refined estimates for the data and computational effort requirements for solving concrete instances of the LWE problem. We apply this refined analysis to suggested parameters for various LWE-based cryptographic schemes from the literature and compare with alternative approaches based on lattice reduction. As a result, we provide new upper bounds for the concrete hardness of these LWE-based schemes. Rather surprisingly, it appears that BKW algorithm outperforms known estimates for lattice reduction algorithms starting in dimension $n \approx 250$ when LWE is reduced to SIS. However, this assumes access to an unbounded number of LWE samples.

6.3.4. Structural Cryptanalysis of McEliece Schemes with Compact Keys

A very popular trend in code-based cryptography is to decrease the public-key size by focusing on subclasses of alternant/Goppa codes which admit a very compact public matrix, typically quasi-cyclic (QC), quasi-dyadic (QD), or quasi-monoidic (QM) matrices. In [5], we show that the very same reason which allows to construct a compact public-key makes the key-recovery problem intrinsically much easier. The gain on the public-key size induces an important security drop, which is as large as the compression factor p on the public-key. The fundamental remark is that from the $k \times n$ public generator matrix of a compact McEliece, one can construct a $k/p \times n/p$ generator matrix which is - from an attacker point of view - as good as the initial public-key. We call this new smaller code the folded code. Any key-recovery attack can be deployed equivalently on this smaller generator matrix. To mount the key-recovery in practice, we also improve the algebraic technique of Faugère, Otmani, Perret and Tillich (FOPT). In particular, we introduce new algebraic equations allowing to include codes defined over any prime field in the scope of our attack. We describe a so-called "structural elimination" which is a new algebraic manipulation which simplifies the key-recovery system. As a proof of concept, we report successful attacks on many cryptographic parameters available in the literature. All the parameters of CFS-signatures based on QD/QM codes that have been proposed can be broken by this approach. In most cases, our attack takes few seconds (the harder case requires less than 2 hours). In the encryption case, the algebraic systems are harder to solve in practice. Still, our attack succeeds against several cryptographic challenges proposed for QD and QM encryption schemes, but there are still some parameters that have been proposed which are out of reach for the methods given here. However, regardless of the key-recovery attack used against the folded code, there is an inherent weakness arising from Goppa codes with QM or QD symmetries. It is possible to derive from the public key a much smaller public key corresponding to the folding of the original QM or QD code, where the reduction factor of the code length is precisely the order of the QM or QD group used for reducing the key size. To summarize, the security of such schemes are not relying on the bigger compact public matrix but on the small folded code which can be efficiently broken in practice with an algebraic attack for a large set of parameters.

6.3.5. A Polynomial-Time Key-Recovery Attack on MQQ Cryptosystems

In [16], we investigate the security of the family of MQQ public key cryptosystems using multivariate quadratic quasigroups (MQQ). These cryptosystems show especially good performance properties. In particular, the

MQQ-SIG signature scheme is the fastest scheme in the ECRYPT benchmarking of cryptographic systems (eBACS). We show that both the signature scheme MQQ-SIG and the encryption scheme MQQ-ENC, although using different types of MQQs, share a common algebraic structure that introduces a weakness in both schemes. We use this weakness to mount a successful polynomial time key-recovery attack that finds an equivalent key using the idea of so-called good keys. In the process we need to solve a MinRank problem that, because of the structure, can be solved in polynomial-time assuming some mild algebraic assumptions. We highlight that our theoretical results work in characteristic 2 which is known to be the most difficult case to address in theory for MinRank attacks and also without any restriction on the number of polynomials removed from the public-key. This was not the case for previous MinRank like-attacks against MQ schemes. From a practical point of view, we are able to break an MQQ-SIG instance of 80 bits security in less than 2 days, and one of the more conservative MQQ-ENC instances of 128 bits security in little bit over 9 days. Altogether, our attack shows that it is very hard to design a secure public key scheme based on an easily invertible MQQ structure.

6.3.6. Algebraic Cryptanalysis of a Quantum Money Scheme The Noise-Free Case

In [14], we investigate the Hidden Subspace Problem (HSP_q) over \mathbb{F}_q which is as follows:

Input : $p_1, \dots, p_m, q_1, \dots, q_m \in \mathbb{F}_q[x_1, \dots, x_n]$ of degree $d \geq 3$ (and $n \leq m \leq 2n$).

Find : a subspace $A \subset \mathbb{F}_q^n$ of dimension $n/2$ (n is even) such that

$$p_i(A) = 0 \quad \forall i \in \{1, \dots, m\} \quad \text{and} \quad q_j(A^\perp) = 0 \quad \forall j \in \{1, \dots, m\},$$

where A^\perp denotes the orthogonal complement of A with respect to the usual scalar product in \mathbb{F}_q .

This problem underlies the security of the first public-key quantum money scheme that is proved to be cryptographically secure under a non quantum but classic hardness assumption. This scheme was proposed by S. Aaronson and P. Christiano at STOC'12. In particular, it depends upon the hardness of HSP_2 . More generally, Aaronson and Christiano left as an open problem to study the security of the scheme for a general field \mathbb{F}_q . We present a randomized polynomial-time algorithm that solves the HSP_q for $q > d$ with success probability $\approx 1 - 1/q$. So, the quantum money scheme extended to \mathbb{F}_q is not secure for big q . Finally, based on experimental results and a structural property of the polynomials that we prove, we conjecture that there is also a randomized polynomial-time algorithm solving the HSP_2 with high probability. To support our theoretical results we also present several experimental results confirming that our algorithms are very efficient in practice. We emphasize that S. Aaronson and P. Christiano proposes a non-noisy and a noisy version of the public-key quantum money scheme. The noisy version of the quantum money scheme remains secure.

6.3.7. Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups

The main practical limitation of the McEliece public-key encryption scheme is probably the size of its key. A famous trend to overcome this issue is to focus on subclasses of alternant/Goppa codes with a non trivial automorphism group. Such codes display then symmetries allowing compact parity-check or generator matrices. For instance, a key-reduction is obtained by taking quasi-cyclic (QC) or quasi-dyadic (QD) alternant/Goppa codes. We show that the use of such symmetric alternant/Goppa codes in cryptography introduces a fundamental weakness. It is indeed possible to reduce the key-recovery on the original symmetric public-code to the key-recovery on a (much) smaller code that has not anymore symmetries. This result [4] is obtained thanks to a new operation on codes called folding that exploits the knowledge of the automorphism group. This operation consists in adding the coordinates of codewords which belong to the same orbit under the action of the automorphism group. The advantage is twofold: the reduction factor can be as large as the size of the orbits, and it preserves a fundamental property: folding the dual of an alternant (resp. Goppa) code provides the dual of an alternant (resp. Goppa) code. A key point is to show that all the existing constructions of alternant/Goppa codes with symmetries follow a common principal of taking codes whose support is globally invariant under the action of affine transformations (by building upon prior works of T. Berger and A. Dür). This enables not only to present a unified view but also to generalize the construction of QC, QD and even

quasi-monoidic (QM) Goppa codes. All in all, our results can be harnessed to boost up any key-recovery attack on McEliece systems based on symmetric alternant or Goppa codes, and in particular algebraic attacks.

6.3.8. Improved Sieving on Algebraic Curves

The best algorithms for discrete logarithms in Jacobians of algebraic curves of small genus are based on index calculus methods coupled with large prime variations. For hyperelliptic curves, relations are obtained by looking for reduced divisors with smooth Mumford representation (Gaudry); for non-hyperelliptic curves it is faster to obtain relations using special linear systems of divisors (Diem, Diem and Kochinke). Recently, Sarkar and Singh have proposed a sieving technique, inspired by an earlier work of Joux and Vitse, to speed up the relation search in the hyperelliptic case. In [20], we give a new description of this technique, and show that this new formulation applies naturally to the non-hyperelliptic case with or without large prime variations. In particular, we obtain a speed-up by a factor approximately 3 for the relation search in Diem and Kochinke's methods.

POSET Team

7. New Results

7.1. Efficient interactive score

We have proposed a solution to the problem of real-time performance for interactive multimedia applications, specifically in the interpretation of interactive multimedia scores. For that, we have proposed a new parallel implementation of interactive scores on a reconfigurable hardware. We take advantage of the parallelism and reliability provided by Field Programmable Gate Arrays (FPGAs) to perform in real-time the hardware representation of scores. The results of the simulations show that our approach allows the system to react instantaneously to user interactions. Moreover, the real-time constraints of the score are satisfied [21].

7.2. Modeling with tile

In [3], [8] it has been observed that musical objects are conveniently modeled by tiles. These modeling experiments have been continued this year by showing, in particular, how both high-level music modeling and low-level signal combination can be modeled by means of tiles [23]. This has been further extended relating classical musical constructs with tile modeling features [34].

7.3. Tiles and inverse semigroups

In [10] it has already been observed that the theory of inverse semigroups⁰ is the adequate mathematical framework to define and study tiles and their languages. In this direction, we have shown that strings, trees and even many types of graphs can be unified into a notion of higher-dimensional strings [24], [35]. Using techniques of partial algebra [4], this notion recovers advanced results on formal languages of graphs of bounded tree-width⁰, which shows the robustness of the approach.

7.4. Reactive programming with tile

The first step towards programming music with tiles is proposed as a Domain-Specific Language : the T-calculus [9]. Further collaboration with Paul Hudak [7] led us to various implementation experiments on top of Haskell [30], [33], [29]. Within the ADT “Tuilage” and S. Archipoff’s PhD thesis in progress, we eventually managed to integrate tile modeling into reactive programming as illustrated, in December 2015, by the first concert of the Idex Arts & Science project “Sound of Algorithm” in collaboration with the musician Edwin Buger.

7.5. Behavioral properties of higher-order programs

In a series of results [28], [27], we have been able to cast to traditional denotational semantics the behavioral properties captured by Monadic Second Order Logic (MSOL) and weak MSOL. The main difficulty was to represent infinitary properties in finitary models. From a foundational point of view, these results exhibit once more the robustness of approaches based on recognizability to capture complex properties of programs. They also make salient the problem of program evaluation in finite models as a milestone towards effective model-checking of higher-order programs.

⁰See [43] for general presentation of inverse semigroup theory, and [45], [44] for graph-based representation of inverse semigroup elements.

⁰See [38] for an up-to-date presentation of the formal language theory of graphs.

7.6. Art & Science project

This year has seen the members of PoSET involved in a number of Art & Science projection, especially some granted by Idex Bordeaux, including but not limited to : *Illumination* created Aurelio Edler-Copes, in partnership with compagnie Eclats, performed in November 2015 at Molière Theater, *Mobiles* and *Le Chant du filament #2* respectively created by Renaud Rubiano and Nicolas Villenave, performed or displayed in November 2015 during FACTS festival, the Art and Science Festival of Bordeaux University, and, *Le son des algorithmes* with Edwin Buger that led to a first musical performance in December 2015.

POSTALE Team

6. New Results

6.1. Parallel light speed labeling: the world's fastest connected component labeling for multicore processors

Participants: Lionel Lacassagne, Laurent Cabaret, Daniel Etiemble.

We have designed a parallel version of the Light Speed Labeling for shared-memory multicore processor. This algorithm outperforms the best algorithm by a factor $\times 10$. We are now working on the design of algorithms for GPU and manycore embedded processor and especially the TSAR architecture of LIP6 laboratory. More information is available at

- TSAR architecture: <https://www-soc.lip6.fr/trac/tsar>
- ALMOS operating system: <https://www-soc.lip6.fr/trac/almos>
- GIET-VM system: <https://www-soc.lip6.fr/trac/giet-vm>

The paper [20] introduces the parallel version of the Light Speed Labeling (LSL) and compares it with the parallel versions of the competitors. A benchmark shows that the parallel Light Speed Labeling is at least $\times 1.9$ faster than all the other algorithms for random images. This factor reach $\times 3.6$ for structured random images. More important, we show that thanks to its run-based processing (segments), LSL is intrinsically more efficient than all pixel-based algorithms.

6.2. Opening Polyhedral Compiler's Black Box

Participants: Lénaïc Bagnères, Oleksandr Zinenko, Stéphane Huot, Cédric Bastoul.

While compilers offer a fair trade-off between productivity and executable performance in single-threaded execution, their optimizations remain fragile when addressing compute-intensive code for parallel architectures with deep memory hierarchies. Moreover, these optimizations operate as black boxes, impenetrable for the user, leaving them with no alternative to time-consuming and error-prone manual optimization in cases where an imprecise cost model or a weak analysis resulted in a bad optimization decision. To address this issue, we propose a technique allowing to automatically translate an arbitrary polyhedral optimization, used internally by loop-level optimization frameworks of several modern compilers, into a sequence of comprehensible syntactic transformations as long as this optimization focuses on scheduling loop iterations. With our approach, we open the black box of the polyhedral frameworks enabling users to examine, refine, replay and even design complex optimizations semi-automatically in partnership with the compiler. [17]

6.3. Automating Resource Selection and Configuration in Inter-clouds through a Software Product Line Method

Participants: Alexandro Ferreira Leite, Vladimir Castro Alves, Genaina Nunes Rodrigues, Claude Tadonki, Christine Eisenbeis, Alba Cristina Alves de Melo.

Nowadays, cloud users face three important problems: (a) choosing one or more appropriate cloud provider(s) to run their application(s), (b) selecting appropriate cloud resources, which implies having enough information about the available resources, including their characteristics and constraints, and (c) configuring the cloud resources. These problems are mostly due to the wide range of resources. These resources usually have distinct dependencies, and they are offered at various clouds' layers. In this complex scenario, the users often have to handle cloud resources and their dependencies manually. This is an error-prone and time-consuming activity, even for skilled cloud users and system administrators. In this context, this paper proposes a software product line engineering (SPLE) method and a tool to deal with these issues. Our SPL-based engineering method enables a declarative and goal-oriented strategy. Furthermore, it allows resource selection and configuration in inter-cloud environments. In our proposal, the cloud users specify their applications and requirements, and our tool automatically selects and configures a suitable computing environment, taking into account temporal and functional dependencies. Experimental results on Amazon EC2 and Google Compute Engine (GCE) show that our approach enables unskilled users to have access to advanced inter-cloud computing configurations, without being concerned with the characteristics of each cloud. [18]

6.4. A Randomized LU-based Solver Using GPU and Intel Xeon Phi Accelerators

Participants: Marc Baboulin, Amal Khabou, Adrien Rémy de Zotti.

We present a fast hybrid solver for dense linear systems based on LU factorization. To achieve good performance, we avoid pivoting by using random butterfly transformations for which we developed efficient implementations on heterogeneous architectures. We used both Graphics Processing Units and Intel Xeon Phi as accelerators. The performance results show that the pre-processing due to randomization is negligible and that the solver outperforms the corresponding routines based on partial pivoting. [16]

6.5. Metaprogramming dense linear algebra solvers. Applications to multi and many-core architectures

Participants: Ian Masliah, Marc Baboulin, Joël Falcou.

The increasing complexity of new parallel architectures has widened the gap between adaptability and efficiency of the codes. As high performance numerical libraries tend to focus more on performance, we wish to address this issue using a C++ library called NT2. By analyzing the properties of the linear algebra domain that can be extracted from numerical libraries and combining them with architectural features, we developed a generic approach to solve dense linear systems on various architectures including CPU and GPU. We have then extended our work with an example of a least squares solver based on semi-normal equations in mixed precision that cannot be found in current libraries. For the automatically generated solvers, we report performance comparisons with state-of-the-art codes, and show that it is possible to obtain a generic code with a high-level interface (similar to MATLAB) which runs either on CPU or GPU without generating a significant overhead. [21] [23]

6.6. Using Random Butterfly Transformations in Parallel Schur Complement-Based Preconditioning

Participants: Marc Baboulin, Aygul Jamal, Masha Sosonkina.

We propose to use a randomization technique based on Random Butterfly Transformations (RBT) in the Algebraic Recursive Multilevel Solver (ARMS) to improve the preconditioning phase in the iterative solution of sparse linear systems. We integrated the RBT technique into the parallel version of ARMS (pARMS). The preliminary experimental results on some matrices from the Davis' collection show an improvement of the convergence and accuracy of the results when compared with existing implementations of the pARMS preconditioner. [15]

6.7. LU Preconditioning for Overdetermined Sparse Least Squares Problems

Participants: Gary Howell, Marc Baboulin.

We investigate how to use an LU factorization with the classical LSQR routine for solving overdetermined sparse least squares problems. Usually L is much better conditioned than A and iterating with L instead of A results in faster convergence. When a runtime test indicates that L is not sufficiently well-conditioned, a partial orthogonalization of L accelerates the convergence. Numerical experiments illustrate the good behavior of our algorithm in terms of storage and convergence. [19]

6.8. Dense Symmetric Indefinite Factorization on GPU Accelerated Architectures

Participants: Marc Baboulin, Jack Dongarra, Adrien Rémy de Zotti, Stanimire Tomov, Ichitaro Yamazaki.

We study the performance of dense symmetric indefinite factorizations (Bunch-Kaufman and Aasen's algorithms) on multicore CPUs with a Graphics Processing Unit (GPU). Though such algorithms are needed in many scientific and engineering simulations, obtaining high performance of the factorization on the GPU is difficult because the pivoting that is required to ensure the numerical stability of the factorization leads to frequent synchronizations and irregular data accesses. As a result, until recently, there has not been any implementation of these algorithms on hybrid CPU/GPU architectures. To improve their performance on the hybrid architecture, we explore different techniques to reduce the expensive communication and synchronization between the CPU and GPU, or on the GPU. We also study the performance of a symmetric indefinite factorization with no pivoting combined with the preprocessing technique based on Random Butterfly Transformations. Though such transformations only have probabilistic results on the numerical stability, they avoid the pivoting and obtain a great performance on the GPU. [14]

6.9. Computing least squares condition numbers on hybrid multicore/GPU systems

Participants: Marc Baboulin, Jack Dongarra, Rémi Lacroix.

We present an efficient computation for least squares conditioning or estimates of it. We propose performance results using new routines on top of the multicore-GPU library MAGMA. This set of routines is based on an efficient computation of the variance-covariance matrix for which, to our knowledge, there is no implementation in current public domain libraries LAPACK and ScaLAPACK. [22]

6.10. Towards a High-Performance Tensor Algebra Package for Accelerators

Participants: Marc Baboulin, Veselin Dobrev, Jack Dongarra, Christopher Earl, Joël Falcou, Azzam Haidar, Ian Karlin, Tzanio Kolev, Ian Masliah, Stanimire Tomov.

Numerous important applications, e.g., high-order FEM simulations, can be expressed through tensors. Examples are computation of FE matrices and SpMV products expressed as generalized tensor contractions. Contractions by the first index can often be represented as tensor index reordering plus gemm, which is a key factor to achieve high-performance. We present ongoing work on the design of a high-performance package in MAGMA for Tensor algebra that includes techniques to organize tensor contractions, data storage, and parametrization related to batched execution of large number of small tensor contractions. We apply auto-tuning and code generation techniques to provide an architecture-aware, user-friendly interface. [24]

PRIVATICS Project-Team

6. New Results

6.1. Surveillance

Participants: Claude Castelluccia, Javier Parra Arnau.

In recent times, we are witnessing an increasing concern by governments and intelligence agencies to deploy mass-surveillance systems that help them fight terrorism. In [40], we conduct a formal analysis of the overall cost of such surveillance systems. Our analysis starts with a fairly-known result in statistics, namely, the false-positive paradox. We propose a quantitative measure of the total cost of a monitoring program, and study a detection system that is designed to minimize it, subject to a constraint in the number of terrorists the agency wishes to capture. In the absence of real, accurate behavioral models, we perform our analysis on the basis of several simple but insightful examples. With these examples, we illustrate the different parameters involved in the design of the detection system, and provide some indicative and representative figures of the cost of the monitoring program.

6.2. Security or privacy ?

Participants: Amrit Kumar, Cédric Lauradoux.

Security softwares such as anti-viruses, IDS or filters help Internet users to protect their privacy from hackers. The cost of this protection is that the users privacy is stripped away by the companies providing these security solutions. Currently, Internet users can choose between no security with the risk of being hacked or use security softwares and lose all personal data to security companies. As a example of this dilemma, we analyze the solution proposed by Google for Safe Browsing in [29] and shows that their privacy policies do not match the reality: Google can perform tracking.

6.3. Users characterization

Participants: Jagdish Acharya, Gergely Acs, Claude Castelluccia.

Prior works have shown that the list of apps installed by a user reveal a lot about user interests and behavior. These works rely on the semantics of the installed apps and show that various user traits could be learnt automatically using off-the-shelf machine-learning techniques. In this work, we focus on the re-identifiability issue and thoroughly study the unicity of smartphone apps on a dataset containing 54,893 Android users collected over a period of 7 months. Our study finds that any 4 apps installed by a user are enough (more than 95% times) for the re-identification of the user in our dataset. As the complete list of installed apps is unique for 99% of the users in our dataset, it can be easily used to track/profile the users by a service such as Twitter that has access to the whole list of installed apps of users. As our analyzed dataset is small as compared to the total population of Android users, we also study how unicity would vary with larger datasets. This work emphasizes the need of better privacy guards against collection, use and release of the list of installed apps.

6.4. Data anonymization

Participants: Claude Castelluccia, Gergely Acs.

Set-valued dataset contains different types of items/values per individual, for example, visited locations, purchased goods, watched movies, or search queries. As it is relatively easy to re-identify individuals in such datasets, their release poses significant privacy threats. Hence, organizations aiming to share such datasets must adhere to personal data regulations. In order to get rid of these regulations and also to benefit from sharing, these datasets should be anonymized before their release. In this paper, we revisit the problem of anonymizing set-valued data. We argue that anonymization techniques targeting traditional k^m -anonymity model, which limits the adversarial background knowledge to at most m items per individual, are impractical for large real-world datasets. Hence, we propose in [25] a probabilistic relaxation of k^m -anonymity and present an anonymization technique to achieve it. This relaxation also improves the utility of the anonymized data. We also demonstrate the effectiveness of our scalable anonymization technique on a real-world location dataset consisting of more than 4 million subscribers of a large European telecom operator. We believe that our technique can be very appealing for practitioners willing to share such large datasets.

6.5. Wi-Fi and privacy

Participants: Jagdish Achara, Mathieu Cunche, Vincent Roca, Celestin Matte.

- **Geolocation spoofing attack** Our work at WiSec 2015 [17] shows how it is possible to manipulate the geolocation information of a single device and how to exploit this information as a side channel to identify the owner of the device on geotagged platforms such as social networks.
- **Extraction of sementical information from Wi-Fi network identifiers** Methods based on text similarity metrics can be used to infer user's interests based on the list of their preferred networks. We present in [23] a method identifying the physical entity (shop, restaurant, company ...) associated to Wi-Fi networks identifiers (SSID).

6.6. Formal and legal issues of privacy

Participants: Thibaud Antignac, Daniel Le Metayer.

- **Privacy by design** Privacy by design will become a legal obligation in the European Community when the Data Protection Regulation eventually gets adopted. However, taking into account privacy requirements in the design of a system is a challenging task. We have proposed an approach based on the specification of privacy architectures and illustrated our formal framework through several case studies. In collaboration with Morpho, we have applied it in the context of biometrics systems. The choice of particular techniques and the role of the components (central server, secure module, terminal, smart card, etc.) in the architecture have a strong impact on the privacy guarantees provided by a biometric system. However, existing proposals were made on a case by case basis, which makes it difficult to compare them and to provide a rationale for the choice of specific options. We have shown that a general framework for the definition of privacy architectures can be used to specify these options and to reason about them in a formal way. In 2015 the results on biometrics were presented at the conferences FM2015 [16] and ISC 2015 [15] (best paper award) and the general approach itself has led to Thibaud Antignac's PhD defense.
- **Verification of privacy properties**

Electric vehicles are an up-and-coming technology that provides significant environmental benefits. A major challenge of these vehicles is their somewhat limited range, requiring the deployment of many charging stations. To effectively deliver electricity to vehicles and guarantee payment, a protocol was developed as part of the ISO 15118 standardization effort. A privacy-preserving variant of this protocol, POPCORN, has been proposed in recent work, claiming to provide significant privacy for the user, while maintaining functionality. We have proposed an approach for the verification of privacy properties of the protocol. We have provided a formal model of the expected privacy properties in the applied Pi-Calculus and used ProVerif to check them. We have identified weaknesses in the protocol in [11] and suggest improvements to address them.
- **Control over personal data**

More than ever the notion of control plays a pivotal and pervasive role in the discourses of privacy and data protection. Privacy scholarship and regulators propose to increase individual control over personal information as the ultimate prescriptive solution to tackle the issues raised by emergent data processing technologies. Conceived as the claim of individuals to determine for themselves when, how, and to what extent information about them is communicated to others, the notion of control is not new. It is often considered as the unique means of empowerment of the data subject. The mechanisms of this empowerment remain however surprisingly vague and understudied. What does it really mean to be in control of one's data in the context of contemporary socio-technical environments and practices? What are the characteristics, purposes and potential limits of such control and how can we guarantee data subjects effective control over their own data? We have carried out an interdisciplinary review of the concept of control to explore such questions in the fields of law and computer science and suggested conditions for the effective application of this concept (see [5]).

- **Accountability** The use of body-worn cameras by police forces around the world is spreading quickly. The resulting mobile and ubiquitous surveillance is often marketed as an instrument for accountability and an effective way of reducing violence. It also involves remarkable potential for intrusion into the privacy of both individuals and police agents. We have studied in [4] the deployment of police body-worn cameras in five countries, investigated their suitability as an accountability tool given the associated privacy threats, and analyzed the societal impact of their deployment as well as the risk of function creep.

6.7. Building blocks

Participant: Marine Minier.

- **Symmetric cryptography** During this year, a fruitful work in collaboration with Céline Blondeau from University of Aalto has appeared in FSE 2015 [8] concerning the equivalence between the key recovery parts of the three attacks (Zero-Correlation, impossible and integral) using the matrix method.

With Thierry Berger, Julien Francq and also Gaël Thomas, we have proposed 2 new lightweight block ciphers : Lilliput and CubeCipher.

Concerning symmetric cryptography, we obtain some results in both sides: on the one hand, we provide 2 new families of lightweight block ciphers: CubeCipher family and Lilliput; on the other hand, we work on the matrix method to simplify the representation of some attacks such as zero-correlation attack, impossible and integral attacks.

We also published the extended version of our Secrypt 2013 paper in the journal Security and Communication Networks [2] concerning the performances on a dedicated platform.

- **Passwords Cracking** Passwords are widely used for user authentication, and will likely remain in use in the foreseeable future, despite several weaknesses. One important weakness is that human-generated passwords are far from being random, which makes them susceptible to guessing attacks. Understanding the adversaries' capabilities for guessing attacks is a fundamental necessity for estimating their impact and advising countermeasures. We develop OMEN [9], a new Markov model-based password cracker that extends ideas proposed by Narayanan and Shmatikov (CCS 2005). The main novelty of our tool is that it generates password candidates according to their occurrence probabilities, i.e., it outputs most likely passwords first. As shown by our extensive experiments, OMEN significantly improves guessing speed over existing proposals. In particular, we compare the performance of OMEN with the Markov mode of John the Ripper, which implements the password indexing function by Narayanan and Shmatikov. OMEN guesses more than 40% of passwords correctly with the first 90 million guesses, while JtR-Markov (for $T = 1$ billion) needs at least eight times as many guesses to reach the same goal, and OMEN guesses more than 80% of passwords correctly at 10 billion guesses, more than all probabilistic password crackers we compared against.

- **Time-memory trade-off** Cryptanalytic time-memory trade-offs (TMTO) are well-known tools available in any security expert toolbox. They have been used to break ciphers such as A5/1, but their efficiency to crack passwords made them even more popular in the security community. While symmetric keys are generated randomly according to a uniform distribution, passwords chosen by users are in practice far from being random, as confirmed by recent leakage of databases. Unfortunately, the technique used to build TMTOs is not appropriate to deal with non-uniform distributions. In [6], we introduce an efficient construction that consists in partitioning the search set into subsets of close densities, and a strategy to explore the TMTOs associated to the subsets based on an interleaved traversal. This approach results in a significant improvement compared to currently used TMTOs. We experimented our approach on a classical problem, namely cracking 7-character NTLM Hash passwords using an alphabet with 34 special characters, which resulted in a $16 \times$ speedup over rainbow tables, which are considered as the most efficient variant of time-memory trade-offs.

PROSECCO Project-Team

7. New Results

7.1. Verification of Security Protocols in the Symbolic Model

Participants: Bruno Blanchet, Miriam Paiola.

The applied pi calculus is a widely used language for modeling security protocols, including as a theoretical basis of **PROVERIF**. However, the seminal paper that describes this language [24] does not come with proofs, and detailed proofs for the results in this paper were never published. This year, Martín Abadi, Bruno Blanchet, and Cédric Fournet finished the detailed proofs of all results of this paper, started last year, and added a new example on a symbolic analog of indifferentiability of hash functions. This work is submitted to a journal.

Previously [37], Bruno Blanchet and Miriam Paiola presented an automatic technique for proving secrecy and authentication properties for security protocols that manipulate lists of unbounded length, for an unbounded number of sessions. That work relies on an extension of Horn clauses, generalized Horn clauses, designed to support unbounded lists, and on a resolution algorithm on these clauses. However, in that previous work, they had to model protocols manually with generalized Horn clauses, which is unpractical. They recently extended the input language of ProVerif to model protocols with lists of unbounded length. They give the formal meaning of this extension, translate it automatically to generalized Horn clauses, and prove that this translation is sound. This work appears as a research report [21].

We implemented several extensions of ProVerif: Bruno Blanchet and Vincent Cheval improved the algorithm for proving observational equivalence between two processes, by merging them into a single biprocess that encodes the two processes. Bruno Blanchet also introduced a new construct `new` $a[x_1, \dots, x_n]$ in ProVerif which allows to specify the arguments x_1, \dots, x_n used in the internal representation of the fresh name a . This extension allows one to tune the precision and speed of the analysis performed by ProVerif. The extended tool is available at <http://proverif.inria.fr>, and deposited to the APP (*Agence pour la Protection des Programmes*).

Stéphanie Delaune, Mark Ryan, and Ben Smyth [42] introduced the idea of swapping data in order to prove observational equivalence. For instance, ballot secrecy in electronic voting is formalized by saying that A voting a and B voting b is observationally equivalent to (indistinguishable from) A voting b and B voting a . Proving such an equivalence typically requires swapping the votes. However, Delaune et al's approach was never proved correct. Bruno Blanchet and Ben Smyth filled this gap by formalizing the approach and providing a detailed soundness proof. They plan to submit this work to a conference.

7.2. Verification of Security Protocols in the Computational model

Participant: Bruno Blanchet.

Bruno Blanchet implemented several extensions of his computational protocol verifier CryptoVerif. In particular, he improved the global dependency analysis, used in order to show that the result of all tests is independent from some random values. He improved the proof of secrecy properties, in particular to prove forward secrecy properties. He also improved the merging of branches of tests, in particular to be able to merge the two branches of `ifbthen` P_1 `else` P_2 even when variables are renamed between P_1 and P_2 . Finally, he added the display of an explanation of why a cryptographic transformation fails, to make the tool easier to use. The extended tool is available at <http://cryptoverif.inria.fr>.

Within the ANR project AnaStaSec, Bruno Blanchet verified an air-ground avionic security protocol (International Civil Aviation Organization (ICAO) Document 9880: Manual on Detailed Technical Specifications for the Aeronautical Telecommunication Network (ATN) using ISO/OSI standards and protocols, Part IV) using CryptoVerif. He proved entity authentication and message authenticity for the main protocol, in the computational model of cryptography, and made comments on some points that should be clarified in the protocol specification. He presented this work at a meeting of the secure dialog service working group of ICAO, in Toulouse, September 2015. The working group was strongly interested by the presentation and welcomed the proposal to apply these modelling and formal verification techniques as part of its validation activities.

7.3. The F* programming language

Participants: Nikhil Swamy [Microsoft Research], Catalin Hritcu, Chantal Keller [LRI], Aseem Rastogi [Univ of Maryland], Antoine Delignat-Lavaud, Simon Forest, Karthikeyan Bhargavan, Cedric Fournet [Microsoft Research], Pierre-Yves Strub [IMDEA], Markulf Kohlweiss [Microsoft Research], Jean Karim Zinzindohoue, Santiago Zanella Beguelin [Microsoft Research, MSR-Inria].

F* is a new higher order, effectful programming language (like ML) designed with program verification in mind. Its type system is based on a core that resembles System F ω (hence the name), but is extended with dependent types, refined monadic effects, refinement types, and higher kinds. Together, these features allow expressing precise and compact specifications for programs, including functional correctness properties. The F* type-checker aims to prove that programs meet their specifications using an automated theorem prover (usually Z3) behind the scenes to discharge proof obligations. Programs written in F* can be translated to OCaml, F#, or JavaScript for execution. We published a paper on the design, implementation, and formal core of F* at POPL 2016. F* is being developed as an open-source project at GitHub: <https://github.com/FStarLang> and the official webpage is at <http://fstar-lang.org>. We released several beta versions of the software this year.

7.4. Micro-Policies and Secure Compilation

Participants: Catalin Hritcu, Arthur Azevedo de Amorim, Zoi Paraskevopoulou, Nikolaos Giannarakis.

Following on from previous work on the *micro-policy* framework, Catalin Hritcu and his collaborators published new work on applications and efficient implementations of micro-policies. They published work on low-level implementations of micro-policies at ASPLOS 2015 [18]. At IEEE S&P, they published a paper how to write formally verified reference monitors using micro-policies [26].

Other than these published works, Hritcu and his colleagues also worked on using micro-policies to enforce secure information flow at the hardware level [25], and a secure compiler for a high-level language that relies on micro-policies to enforce programming language abstractions [45].

7.5. Dependable Property-Based Testing

Participants: Catalin Hritcu, Zoi Paraskevopoulou.

Catalin Hritcu and his student, Zoi Paraskevopoulou, worked on a methodology for formally verified property-based testing and implemented it as a foundational verification framework for QuickChick, a port of QuickCheck to Coq. This work was published at ITP 2015 [19]. Catalin Hritcu also worked with a number of co-authors on a new technique for creating random generators for property-based testing. This work is currently under submission [46].

7.6. Attacks and Proofs for Transport Layer Security

Participants: Benjamin Beurdouche, Karthikeyan Bhargavan, Antoine Delignat-Lavaud, Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Alfredo Pironti, Pierre-Yves Strub [IMDEA], Jean Karim Zinzindohoue.

As a countermeasure to our earlier work on the triple handshake attack, we proposed a TLS extension called *session hash* which has now been published as an Internet standard (IETF RFC 7627). We also formally analyzed various protocols such as TLS, IKE, and SSH for key synchronization and triple handshake attacks, and proved that our session hash countermeasure prevents such attacks on TLS. This work appeared at NDSS 2015 [15].

We discovered and reported an important class of *state machine attacks* on implementations of the Transport Layer Security (TLS) protocol. These attacks appear when TLS implementations incorrectly accept messages which are forbidden by the TLS state machine. We built a test framework for such attacks and analyzed a number of open source implementations. Our analysis uncovered critical vulnerabilities such as the SKIP attack on Java and the FREAK attack on almost all mainstream web browsers. The research results were published at IEEE S&P where our paper won a distinguished paper award [14]. Our work also led to security updates and CVEs for many web browsers, TLS libraries, and web servers.

Along with colleagues at several other institutions, we discovered the Logjam vulnerability on protocols that still support weak Diffie-Hellman groups in their key exchange. We showed that the attack could be used for online and offline attacks on real-world TLS clients and servers. We also showed how the vulnerability could weaken the security of IPsec and SSH connections. Our research led to widespread changes to the configurations of web servers, mail servers, web browsers, and TLS libraries. The research was published at ACM CCS 2015 [12] where it won a Best Paper award.

Antoine Delignat-Lavaud showed how the unsafe sharing of certificates across multiple HTTPS websites could be exploited to fully compromise the same origin policy for websites, using a vulnerability called *virtual host confusion*. A research paper on these attacks appeared at WWW 2015 [17].

7.7. Privacy, Electronic Voting, and Auctions

Participants: Benjamin Smyth [correspondant], Elizabeth Quaglia.

Benjamin Smyth worked on a formal analysis of privacy in Direct Anonymous Attestation schemes [50]. He also showed how to verify commitment protocols in ProVerif without False attacks [39].

Apart from these published works, Benjamin Smyth and Elizabeth Quaglia worked on formal security analyses of electronic auction schemes based on existing models for electronic voting [48]. Benjamin Smyth worked on developing new formal definitions for secrecy and independence in election schemes [51], and on applying such definitions to the security analysis of real-world voting protocols such as Helios and JCJ [49].

7.8. Computationally Complete Symbolic Attacker Models

Participants: Gergei Bana, Hubert Comon-Lundh [ENS Cachan], Rohit Chadha [University of Missouri].

In previous work, Bana and Comon-Lundh proposed a new approach to computational verification of cryptographic protocols, by defining a *computationally complete* symbolic attacker, so that a symbolic proof against this attacker can be shown to imply a computational proof of security [27], [28].

Following on from this work, Bana and Chadha fully developed the core parts of the computationally complete symbolic attacker based on indistinguishability. This covers both trace properties and equivalence properties and can be proved partially complete. They evaluated their method by applying it to several classic protocols. This work is currently under submission.

Bana, Comon-Lundh, and Koutsos also worked on a decision procedure for the computationally complete symbolic attacker based on indistinguishability.

SECRET Project-Team

6. New Results

6.1. Symmetric cryptology

Participants: Anne Canteaut, Pascale Charpin, Sébastien Duval, Virginie Lallemand, Gaëtan Leurent, Nicky Mouha, María Naya Plasencia, Joëlle Roué, Yann Rotella.

6.1.1. Block ciphers

Most of our work on block ciphers is related to an ANR Project named BLOC. Our recent results mainly concern either the analysis and design of lightweight block ciphers.

Recent results:

- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalizes the so-called α -reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [15], [60].
- Formalization and generic improvements of impossible differential cryptanalysis: our work provides a general framework for impossible differential cryptanalysis including a generic complexity analysis of the optimal attack [36].
- Cryptanalysis of several recently proposed block ciphers which offer an optimal resistance against side-channel attacks in the sense that the cost of Boolean masking is minimized. This includes an attack against Zorro and its variants [39], and an attack against Picaro in the related-key model [44].
- Cryptanalysis of Feistel constructions with secret Sboxes [42].
- Study of the security of the Even-Mansour construction in the multi-key setting [56].

6.1.2. Authenticated encryption

A limitation of all classical block ciphers is that they aim at protecting confidentiality only, while most applications need both encryption and authentication. These two functionalities are provided by using a block cipher like the AES together with an appropriate mode of operation. However, it appears that the most widely-used mode of operation for authenticated encryption, AES-GCM, is not very efficient for high-speed networks. Also, the security of the GCM mode completely collapses when an IV is reused. These severe drawbacks have then motivated an international competition named CAESAR, partly supported by the NIST, which has been recently launched in order to define some new authenticated encryption schemes⁰. Our work related to this competition is then two-fold: G. Leurent and N. Mouha have participated to the design of some CAESAR candidates; Also, the project-team is involved in a national cryptanalytic effort led by the BRUTUS project funded by the ANR.

Recent results:

- Design of new authenticated encryption schemes submitted to the CAESAR competition: SCREAM v3.0 [72] and PRIMATES 2[58]
- Cryptanalysis of the CAESAR candidates: collision attacks [49] against several candidates including AEZ and Marble, attack against LAC [53].

6.1.3. Stream ciphers

Stream ciphers provide an alternative to block-cipher-based encryption schemes. They are especially well-suited in applications which require either extremely fast encryption or a very low-cost hardware implementation.

⁰<http://competitions.cr.yp.to/caesar.html>

Recent results:

- Cryptanalysis of the recently proposed lightweight stream cipher Sprout [52], [71].
- New types of correlation attacks against filter generators exploiting the approximation of the filtering function composed with non-bijective monomial mappings [63], [87].
- Design of encryption schemes for efficient homomorphic-ciphertext compression: in order to avoid the (extremely) high expansion rate of homomorphic encryption, a solution consists in transmitting to the server the ciphertext c obtained by encrypting m with a symmetric scheme (the corresponding secret key encrypted by the homomorphic cipher is also transmitted). The server then needs to compute m encrypted with the homomorphic scheme from c , i.e. the server needs to homomorphically evaluate the decryption circuit of the symmetric cipher. A. Canteaut, M. Naya-Plasencia together with their coauthors have investigated the constraints on the symmetric cipher imposed by this application and they have proposed some solutions based on additive IV-based stream ciphers [78].

6.1.4. Hash functions and MACS

The international research effort related to the selection of the new hash function standard SHA-3 has led to many important results and to a better understanding of the security offered by hash functions. However, hash functions are used in a huge number of applications with different security requirements, and also form the building-blocks of some other primitives, like MACs. In this context, we have investigated the security of some of these constructions, in order to determine whether some particular constructions for hash functions may affect the security of the associated MACs.

Recent results:

- Improved generic attacks against hash-based MAC [30], [31]
- Cryptanalysis of 7 (out of 8) rounds of the Chaskey MAC [32]. This work has led the designers of Chaskey to increase the number of rounds [80].
- Attack against the XOR of two hash functions, using complex structures build from collisions [54]. This work by G. Leurent and L. Wang shows that, surprisingly, the construction $H_1(M) \oplus H_2(M)$ with common hash functions H_1 and H_2 (e.g. SHA-256 and BLAKE-256) is actually be less secure than each function on their own.

6.1.5. Security of Internet protocols

Hash functions are used to in key-exchange protocols such as TLS, IKE and SSH, to verify the integrity of the exchange. Most practitioners believe that the hash function only need to resist preimage attacks for this use. However, K. Bhargavan and G. Leurent have shown that collisions in the hash function are sufficient to break the integrity of these protocols, and to impersonate some of the parties [41]. Since many protocols still allow the use of MD5 or SHA-1 (for which collision attacks are known), this result in some practical attacks, and extends the real-world impact of the collision attacks against MD5 and SHA-1. This work has already influenced the latest TLS 1.3 draft, and the main TLS libraries are removing support of MD5 signatures

6.1.6. Cryptographic properties and construction of appropriate building blocks

The construction of building blocks which guarantee a high resistance against the known attacks is a major topic within our project-team, for stream ciphers, block ciphers and hash functions. The use of such optimal objects actually leads to some mathematical structures which may be at the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, characterizing the structures of the building blocks which are optimal regarding to some attacks is very important for finding appropriate constructions and also for determining whether the underlying structure induces some weaknesses or not. For these reasons, we have investigated several families of filtering functions and of S-boxes which are well-suited for their cryptographic properties or for their implementation characteristics.

Recent results:

- Definition of an extended criterion for estimating the resistance of a block cipher to differential attacks. This work emphasizes the role played by the affine permutation of the set of 8-bit words which follows the inverse function in the AES [45], [25], [26], [64], [24] (see Section 5.1.1).
- Construction of new Sboxes for lightweight ciphers: A. Canteaut, S. Duval and G. Leurent have investigated several constructions for obtaining good cryptographic Sboxes (especially 8-bit Sboxes) with a low implementation cost [43], [62], [84].
- P. Charpin, together with S. Mesnager and S. Sarkar, has provided a rigorous study of involutions over the finite field of order 2^n which are relevant primitives for cryptographic designs [47]. Most notably, they have focused on the class of involutions defined by Dickson polynomials [70], [79].

6.2. Code-based cryptography

Participants: Rodolfo Canto Torres, Julia Chaulet, Adrien Hauteville, Irene Márquez Corbella, Aurélie Phezzo, Nicolas Sendrier, Jean-Pierre Tillich.

The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Some of the main issues in this field are

- security analysis, implementation and practicality of existing solutions,
- reducing the key size, *e.g.*, by using rank metric instead of Hamming metric, or by using particular families of codes,
- addressing new functionalities, like hashing or symmetric encryption.

Recent results:

- Structural attacks against some variants of the McEliece cryptosystem based on subclasses of alternant/Goppa codes which admit a very compact public matrix, typically quasi-cyclic, quasi-dyadic, or quasi-monoidic matrices [20]. This result is obtained thanks to a new operation on codes called folding that exploits the knowledge of the automorphism group of the code [19].
- Cryptanalysis of a variant of McEliece cryptosystem based on polar codes [40], [59].
- Cryptanalysis of a code-based encryption scheme proposed by Baldi *et al.* in the *Journal of Cryptology* [48].
- Cryptanalysis of a code-based signature scheme proposed at PQCrypto 2013 by Baldi *et al.* [57].
- Improved algorithm for decoding in the rank metric when some additional information about the targeted codeword is provided [51]; this algorithm used together with a folding technique leads to a feasible attack on the LRPC cryptosystem.
- Design on a new code-based stream cipher, named RankSynd, variant of Synd for the rank metric [50].
- In-depth analysis of the complexity of generic decoding algorithms for linear codes [37]. Most notably, R. Canto Torres and N. Sendrier have investigated the information-set decoding algorithms applied to the case where the number of errors is sub-linear in the code length [46]. This situation appears in the analysis of the McEliece based in quasi-cyclic Moderate Density Parity Check (MDPC) codes.

6.3. Quantum Information

Participants: Kaushik Chakraborty, André Chailloux, Anthony Leverrier, Jean-Pierre Tillich.

6.3.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It is also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time.

Recent results:

- A. Leverrier and JP. Tillich, together with G. Zémor, proposed a new class of quantum LDPC codes, “Quantum expander codes”, which feature a simple and very efficient decoding algorithm which can correct arbitrary patterns of errors of size scaling as the square-root of the length of the code. These are the first codes with constant rate for which such an efficient decoding algorithm is known (see Section 5.1.3) [55], [35], [73].
- Error analysis for Boson Sampling, a simplified model for quantum computation [21]

6.3.2. Quantum cryptography

A recent approach to cryptography takes into account that all interactions occur in a physical world described by the laws of quantum physics. These laws put severe constraints on what an adversary can achieve, and allow for instance to design provably secure key distribution protocols. We study such protocols as well as more general cryptographic primitives such as coin flipping with security properties based on quantum theory.

Recent results:

- A. Leverrier gave the first composable security proof for a continuous-variable quantum key distribution protocol with coherent states [22]. This essentially completes the security analysis of continuous-variable protocols with coherent states, which are by far the most practical protocols relying on continuous variables.
- A. Leverrier and E. Diamanti reviewed the state-of-the-art concerning quantum key distribution with continuous variables [18].
- A. Leverrier and M. Tomamichel gave the most complete security proof of the BB84 protocol to date, including all finite-size effects and a full description of the protocol [89].
- K. Chakraborty and A. Leverrier studied a general family of quantum protocols for position verification and present a new class of attacks based on the Clifford hierarchy that outperform previously known attacks [17].

6.3.3. Quantum correlations and nonlocality

Since the seminal work from Bell in the 60’s, it has been known that classical correlations obtained via shared randomness cannot reproduce all the correlations obtained by measuring entangled quantum systems. This impossibility is for instance witnessed by the violation of a Bell inequality and is known under the name of “Quantum Nonlocality”. In addition to its numerous applications for quantum cryptography, the study of quantum nonlocality and quantum games has become a central topic in quantum information theory, with the hope of bringing new insights to our understanding of quantum theory.

Recent results:

- Development of a general framework for the study of quantum correlations with combinatorial tools [14]

6.3.4. Relativistic cryptography

(see Section 5.1.2).

6.3.5. *Quantum cryptanalysis of symmetric primitives*

Symmetric cryptography seems at first sight much less affected in the post-quantum world than asymmetric cryptography: its main known threat is Grover's algorithm, which allows for an exhaustive key search in the square root of the normal complexity. For this reason, it is usually believed that doubling key lengths suffices to maintain an equivalent security in the post-quantum world. However, a lot of work is certainly required in the field of symmetric cryptography in order to "quantize" the classical families of attacks in an optimized way. G. Leurent, A. Leverrier and M. Naya Plasencia have recently started working in this area in collaboration with M. Kaplan, especially on differential cryptanalysis. Some preliminary results show that counter-intuitive and surprising cases appear: in general, it is not sufficient to consider the best classical attacks and try to "quantize" them if one wants to find the best post-quantum attack [34], [85].

6.4. Reverse-engineering of communication systems

Participants: Nicolas Sendrier, Jean-Pierre Tillich, Audrey Tixier.

Our activity within this domain, whose first aim is to establish the scientific and technical foundations of a discipline which does not exist yet at an academic level, has been supported by some industrial contracts driven by the Ministry of Defense.

Recent results:

- Efficient algorithm for recovering the block interleaver and the convolutional code when several noisy interleaver codewords are given [76], [13].

SPADES Project-Team

6. New Results

6.1. Components and contracts

Participants: Sophie Quinton, Jean-Bernard Stefani.

6.1.1. Multi-viewpoint contracts for the negotiation of embedded software updates

In the context of the CCC project (<http://ccc-project.org/>) we address the issue of change after deployment in safety-critical embedded system applications. Our goal is to substitute lab-based verification with in-field formal analysis to determine whether an update may be safely applied. This is challenging because it requires an automated process able to handle multiple viewpoints such as functional correctness, timing, etc. For this purpose, we propose an original methodology for contract-based negotiation of software updates. The use of contracts allows us to cleanly split the verification effort between the lab and the field. In addition, we show how to rely on existing viewpoint-specific methods for update negotiation. We have started validating our approach on a concrete example inspired by the automotive domain in collaboration with our German partners from TU Braunschweig.

6.1.2. Location Graphs

The design of configurable systems can be streamlined and made more systematic by adopting a component-based structure, as demonstrated with the FRACTAL component model [2]. However, the formal foundations for configurable component-based systems, featuring higher-order capabilities where components can be dynamically instantiated and passivated, and non-hierarchical structures where components can be contained in different composites at the same time, are still an open topic. We have recently introduced the location graph model [88], where components are understood as graphs of locations hosting higher-order processes, and where component structures can be arbitrary graphs.

We have continued the development of the location graph model and extended it in several directions. First we have introduced basic capabilities and predicate parameters in the model to allow for different forms of architectural invariants, such as different forms of encapsulation, to be maintained even in presence of dynamic graph modifications. Second, we have started developing the premises of a refinement theory for location graphs, showing in particular how one could refine a location process into a whole graph. Finally, we have shown how to handle heterogeneous forms of composition in the same location graph, turning each location into a composition operator. This work has not yet been published.

6.2. Real-Time multicore programming

Participants: Vagelis Bebelis, Adnan Bouakaz, Pascal Fradet, Alain Girault, Gregor Goessler, Xavier Nicollin, Jean-Bernard Stefani.

6.2.1. A time predictable programming language for multicores

Time predictability (PRET) is a topic that emerged in 2007 as a solution to the ever increasing unpredictability of today's embedded processors, which results from features such as multi-level caches or deep pipelines [59]. For many real-time systems, it is mandatory to compute a strict bound on the program's execution time. Yet, in general, computing a tight bound is extremely difficult [92]. The rationale of PRET is to simplify both the programming language and the execution platform to allow more precise execution times to be easily computed [38].

Following our past results on the PRET-C programming language [36], we have proposed a time predictable synchronous programming language for multicores, called FOREC. It extends C with a small set of ESTEREL-like synchronous primitives to express concurrency, interaction with the environment, looping, and a synchronization barrier [93] (like the pause statement in ESTEREL). FOREC threads communicate with each other via shared variables, the values of which are *combined* at the end of each tick to maintain deterministic execution. FOREC is compiled into threads that are then statically scheduled for a target multicore chip. Our WCET analysis takes into account the access to the shared TDMA bus and the necessary administration for the shared variables. We achieve a very precise WCET (the over-approximation being less than 2%) thanks to a reachable space exploration of the threads' states.

Recent results have addressed the semantics, the compiler, and the experiments. In particular, we have sought to provide several combine policies for shared variables, in a way similar as concurrent revisions [49].

This work has been conducted within the RIPPES associated team.

6.2.2. Modular distribution of synchronous programs

Synchronous programming languages describe functionally centralized systems, where every value, input, output, or function is always directly available for every operation. However, most embedded systems are nowadays composed of several computing resources. The aim of this work is to provide a language-oriented solution to describe *functionally distributed reactive systems*. This research started within the Inria large scale action SYNCHRONICS and is a joint work with Marc Pouzet (ENS, PARKAS team from Rocquencourt) and Gwenaël Delaval (UGA, CTRL-A team from Grenoble).

We are working on defining a *fully-conservative* extension of a synchronous data-flow programming language (the HEPTAGON language, inspired from LUCID SYNCHRONE [51]). The extension, by means of *annotations* adds *abstract location parameters* to functions, and *communications* of values between locations. At deployment, every abstract location is assigned an actual one; this yields an executable for each actual computing resource. Compared to the PhD of Gwenaël Delaval [56], [57], the goal here is to achieve *modular* distribution even in the presence of non-static clocks, *i.e.*, clocks defined according to the value of inputs.

By *fully-conservative*, we have three aims in mind:

1. A non-annotated (*i.e.*, centralized) program will be compiled exactly as before;
2. An annotated program eventually deployed onto only one computing location will behave exactly as its centralized counterpart;
3. The input-output semantics of a distributed program is the same as its centralized counterpart.

By *modular*, we mean that we want to compile each function of the program into a single function capable of running on any computing location. At deployment, the program of each location may be optimized (by simple Boolean-constant-propagation, dead-code and unused-variable elimination), yielding different optimized code for each computing location.

We have formalized the type-system for inferring the location of each variable and computation. In the presence of local clocks, added information is computed from the existing clock-calculus and the location-calculus, to infer necessary communication of clocks between location. The overall structure of the new compiler is defined, including new algorithms for deployment (and code optimization), achieving the three aims detailed above.

6.2.3. Analysis and scheduling of parametric dataflow models

Recent data-flow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (*e.g.*, Kahn Process Networks or the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking *liveness* (*i.e.*, no part of the system will deadlock) and *boundedness* (*i.e.*, the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems.

Recently, we have introduced the *Schedulable Parametric Data-Flow* (SPDF) MoC for dynamic streaming applications [62], which extends the standard dataflow model by allowing rates to be parametric, and the *Boolean Parametric Data Flow* (BPDF) MoC [42], [41] which combines integer parameters (to express dynamic rates) and boolean parameters (to express the activation and deactivation of communication channels). High dynamicity is provided by integer parameters which can change at each basic iteration and boolean parameters which can change even within the iteration. We have presented static analyses that ensure the liveness and the boundedness of BPDF graphs.

This year, we have focused on the *symbolic* analysis of parametric data-flow graphs. This work has been conducted within the RIPPEES associated team.

6.2.4. *Synthesis of switching controllers using approximately bisimilar multiscale abstractions*

The use of discrete abstractions for continuous dynamics has become standard in hybrid systems design (see e.g., [90] and the references therein). The main advantage of this approach is that it offers the possibility to leverage controller synthesis techniques developed in the areas of supervisory control of discrete-event systems [83]. The first attempts to compute discrete abstractions for hybrid systems were based on traditional systems behavioral relationships such as simulation or bisimulation, initially proposed for discrete systems most notably in the area of formal methods. These notions require inclusion or equivalence of observed behaviors which is often too restrictive when dealing with systems observed over metric spaces. For such systems, a more natural abstraction requirement is to ask for closeness of observed behaviors. This leads to the notions of approximate simulation and bisimulation introduced in [63].

These approaches are based on sampling of time and space where the sampling parameters must satisfy some relation in order to obtain abstractions of a prescribed precision. In particular, the smaller the time sampling parameter, the finer the lattice used for approximating the state-space; this may result in abstractions with a very large number of states when the sampling period is small. However, there are a number of applications where sampling has to be fast; though this is generally necessary only on a small part of the state-space. We have been exploring two approaches to overcome this state-space explosion.

We are currently investigating an approach using mode sequences of given length as symbolic states for our abstractions. By using mode sequences of variable length we are able to adapt the granularity of our abstraction to the dynamics of the system, so as to automatically trade off precision against controllability of the abstract states.

6.2.5. *Typical Worst-Case Analysis of real-time systems*

We focus on the problem of computing tight deadline miss models for real-time systems, which bound the number of potential deadline misses in a given sequence of activations of a task. In practical applications, such guarantees are often sufficient because many systems are in fact not hard real-time. Our major contribution this year is a general formulation of that problem in the context of systems where some tasks occasionally experience sporadic overload [26]. Based on this new formulation, we present an algorithm that can take into account fine-grained effects of overload at the input of different tasks when computing deadline miss bounds. We show in experiments with synthetic as well as industrial data that our algorithm produces bounds that are much tighter than in previous work, in sufficiently short time. In addition, we improve, in the preemptive case, the criterion proposed in [71] for establishing how much overload can be tolerated in a time window while still guaranteeing absence of deadline misses: our new criterion is a necessary and sufficient condition (as opposed to the sufficient condition of [71]) and therefore yields better results.

In parallel, we have developed an extension of sensitivity analysis for budgeting in the design of weakly-hard real-time systems. During design, it often happens that some parts of a task set are fully specified while other parameters, e.g. regarding recovery or monitoring tasks, will be available only much later. In such cases, sensitivity analysis can help anticipate how these missing parameters can influence the behavior of the whole system so that a resource budget can be allocated to them. It is, however, sufficient in many application contexts to budget these tasks in order to preserve weakly-hard rather than hard guarantees. We have thus developed an extension of sensitivity analysis for deriving task budgets for systems with hard and weakly-hard requirements.

We currently validate our approach on synthetic test cases and a realistic case study given by our partner Thales.

6.3. Language Based Fault-Tolerance

Participants: Dmitry Burlyaev, Pascal Fradet, Alain Girault, Yoann Geoffroy, Gregor Goessler, Jean-Bernard Stefani, Atena Abdi, Ismail Assayad.

6.3.1. Fault Ascription in Concurrent Systems

The failure of one component may entail a cascade of failures in other components; several components may also fail independently. In such cases, elucidating the exact scenario that led to the failure is a complex and tedious task that requires significant expertise.

The notion of causality (*did an event e cause an event e' ?*) has been studied in many disciplines, including philosophy, logic, statistics, and law. The definitions of causality studied in these disciplines usually amount to variants of the counterfactual test “ e is a cause of e' if both e and e' have occurred, and in a world that is as close as possible to the actual world but where e does not occur, e' does not occur either”. In computer science, almost all definitions of logical causality — including the landmark definition of [70] and its derivatives — rely on a causal model that may not be known, for instance in presence of black-box components. For such systems, we have been developing a framework for blaming that helps us establish the causal relationship between component failures and system failures, given an observed system execution trace. The analysis is based on a formalization of counterfactual reasoning [7].

We have instantiated our approach to a synchronous data flow framework defined by a subset of the LUSTRE [69] language, and implemented the analysis in LOCA (see Section 5.2).

In [25] we have shown that we can improve precision of the analysis if (1) we can emulate execution of components instead of relying on their specifications, and (2) take into consideration input/output dependencies between components to avoid blaming components for faults induced by other components. We have demonstrated the utility of the extended analysis with a case study for a closed-loop patient-controlled analgesia system.

We have further proposed in [23] a general semantic framework for fault ascription. Our framework relies on configuration structures to handle concurrent systems, partial and distributed observations in a uniform way. It defines basic conditions for a counterfactual analysis of necessary and sufficient causes, and it presents a refined analysis that conforms to our basic conditions while avoiding various infelicities.

6.3.2. Tradeoff exploration between energy consumption and execution time

We have continued our work on multi-criteria scheduling, in two directions. First, in the context of dynamic applications that are launched and terminated on an embedded homogeneous multi-core chip, under execution time and energy consumption constraints, we have proposed a two layer adaptive scheduling method. In the first layer, each application (represented as a DAG of tasks) is scheduled statically on subsets of cores: 2 cores, 3 cores, 4 cores, and so on. For each size of these sets (2, 3, 4, ...), there may be only one topology or several topologies. For instance, for 2 or 3 cores there is only one topology (a “line”), while for 4 cores there are three distinct topologies (“line”, “square”, and “T shape”). Moreover, for each topology, we generate statically several schedules, each one subject to a different total energy consumption constraint, and consequently with a different Worst-Case Reaction Time (WCRT). Coping with the energy consumption constraints is achieved thanks to Dynamic Frequency and Voltage Scaling (DVFS). In the second layer, we use these pre-generated static schedules to reconfigure dynamically the applications running on the multi-core each time a new application is launched or an existing one is stopped. The goal of the second layer is to perform a dynamic global optimization of the configuration, such that each running application meets a pre-defined quality-of-service constraint (translated into an upper bound on its WCRT) and such that the total energy consumption be minimized. For this, we (1) allocate a sufficient number of cores to each active application, (2) allocate the unassigned cores to the applications yielding the largest gain in energy, and (3) choose for each application the best topology for its subset of cores (*i.e.*, better than the by default “line” topology). This is a joint work with Ismail Assayad (U. Casablanca, Morocco) who visited the team in September 2015.

Second, in the context of a static application (again represented a DAG of tasks) running on an homogeneous multi-core chip, we have worked on the static scheduling minimizing the WCRT of the application under the multiple constraints that the reliability, the power consumption, and the temperature remain below some given threshold. There are multiple difficulties: (1) the reliability is not an invariant measure w.r.t. time, which makes it impossible to use backtrack-free scheduling algorithms such as list scheduling [37]; to overcome this, we adopt instead the Global System Failure Rate (GSFR) as a measure of the system's reliability that is invariant with time [64]; (2) keeping the power consumption under a given threshold requires to lower the voltage and frequency, but this has a negative impact both on the WCRT and on the GSFR; keeping the GSFR below a given threshold requires to replicate the tasks on multiple cores, but this has a negative impact both on the WCRT, on the power consumption, and on the temperature; (3) keeping the temperature below a given threshold is even more difficult because the temperature continues to increase even after the activity stops, so each scheduling decision must be assessed not based on the current state of the chip (*i.e.*, the temperature of each core) but on the state of the chip at the end of the candidate task, and cooling slacks must be inserted. This is a joint work with Atena Abdi (Amirkabir U., Iran) who is a PhD visitor in the team.

6.3.3. Automatic transformations for fault tolerant circuits

In the past years, we have studied the implementation of specific fault tolerance techniques in real-time embedded systems using program transformation [1]. We are now investigating the use of automatic transformations to ensure fault-tolerance properties in digital circuits. To this aim, we consider program transformations for hardware description languages (HDL). We consider both single-event upsets (SEU) and single-event transients (SET) and fault models of the form “*at most 1 SEU or SET within n clock cycles*”.

We have proposed novel fault-tolerance transformations based on time-redundancy. In particular, we have presented a transformation using double-time redundancy (DTR) coupled with micro-checkpointing, rollback and a speedup mode [19]. The approach is capable to mask any SET every 10 cycles and keeps the same input/output behavior regardless error occurrences. Usually transparent masking requires triple redundancy and voting. Experimental results on the ITC'99 benchmark suite indicate that the hardware overhead of DTR is 2.7 to 6.1 times smaller than full TMR with a double loss in throughput. The method does not require any specific hardware support and is an interesting alternative to Triple Modular Redundancy (TMR) for logic intensive designs.

We have also designed a transformation that allows the circuit to change its level of time-redundancy. This feature allows the circuit to dynamically and temporarily lower (resp. increase) fault-tolerance and speed up (resp. slow down) its computation without interruption [20]. The motivations for such changes can be based on the current radiation environment or the processing of critical data. When hardware size is limited and fault-tolerance is only occasionally needed, that scheme is a better choice than static TMR, which involves a constant high area overhead

These time redundancy transformations (DTR and adaptive fault-tolerance) have been patented [50]

We have described how to formally certify fault-tolerant transformations using the COQ proof assistant [53] (see Section 5.3). The transformations are described on a simple gate-level hardware description language LDDL (Low-level Dependent Description Language). This combinator language is equipped with dependent types and ensures that circuits are well-formed by construction (gates correctly plugged, no dangling wires, no combinational loops, ...). Fault-models are specified in the operational semantics of the language. The main theorem states that, for any circuit, for any input stream and for any SET allowed by the fault-model, its transformed version produces a correct output [18]. The primary motivation of this work was to certify DTR whose intricacy requested a formal proof to make sure that no single-point of failure existed. We have first applied this approach to the correctness proofs of TMR, TTR (Triple Time Redundancy) and finally DTR.

This research is part of Dmitry Burlyayev's PhD thesis [11] defended in November 2015.

6.3.4. A formal approach for the synthesis and implementation of fault-tolerant embedded systems

We have been working for several years on the usage of discrete controller synthesis (DCS) [83] to provide the automated addition of fault-tolerance in embedded systems with formal guarantees [65]. The first key idea is that the initial system model (usually an LTS) includes both the expected behaviors, the unexpected ones (that is, the failures), and the reconfigurations (typically repair actions). The second key idea is that the failures are modeled as *uncontrollable* events. Then, thanks to an exhaustive state space traversal, DCS is able to generate a *controller* that will prevent the system from entering a “bad” state (*e.g.*, a configuration of the system where a task is active on a faulty processor). From the point of view of fault-tolerance, this approach combines the advantages of static guarantees with that of dynamic reconfiguration (hence without the penalty of static redundancy).

Through this new work, we have demonstrated the feasibility of a complete workflow to synthesize and implement correct-by-construction fault tolerant distributed embedded systems consisting of real-time periodic tasks [24]. Correct-by-construction is provided by the use of DCS, which allows us to guarantee that the synthesized controlled system guarantees the functionality of its tasks even in the presence of processor failures. For this step, our workflow uses the HEPTAGON domain specific language [58] and the SIGALI DCS tool [79]. The correct implementation of the resulting distributed system is a challenge, all the more since the controller itself must be tolerant to the processor failures. We achieve this step thanks to the libDGALS real-time library [89] (1) to generate the glue code that will migrate the tasks upon processor failures, maintaining their internal state through migration, and (2) to make the synthesized controller itself fault-tolerant.

SPECFUN Project-Team

6. New Results

6.1. Integration of rational functions

Periods of rational integrals are specific integrals, with respect to one or several variables, whose integrand is a rational function and whose domain of integration is closed. This particular class of integrals contains large families of functions naturally occurring in combinatorics and statistical physics, such as diagonals, constant terms and positive part of rational functions. Periods involving one parameter are classically known to satisfy *Picard-Fuchs equations*, a special type of linear differential equations with a very rich analytic and arithmetic structure. As for other special-function manipulations, handling periods through those differential equations is a good way to actually compute them, and this was the topic of Pierre Lairez' PhD thesis defended in 2014 [53] and awarded the "Ecole Polytechnique thesis prize" in 2015.

Computing multivariate integrals is one speciality of the team and our algorithms are known to treat much more general integrals than just periods of rational integrals. However, integration is still slow in practice when the number of variables goes increasing. By looking at periods of rational functions, the hope is to obtain relevant complexity bounds and faster algorithms.

The goal of reaching relevant theoretical complexity bounds had been reached in 2013 [31] but a practically fast algorithm was still missing. This year, we described a new algorithm which is efficient in practice [4], though its complexity is not known. This algorithm allows to compute quickly integrals that are too big to be computed with previous algorithms. As a challenging benchmark, we computed 210 integrals given by Batyrev and Kreuzer in their work on Calabi–Yau varieties. This achievement gave strong visibility to the paper and allowed a quick dissemination of the implementation, which is provided in Magma under a CeCILL B license. The algorithm is now used on a regular basis by several teams. We know of:

- Tom Coates' team (Dpt. of Mathematics, Imperial College, London, UK), which uses the software in their work about mirror symmetry and classification of Fano varieties;
- Duco van Straten (Institute of Mathematics, University of Mainz, Germany), who uses the software in his work in algebraic geometry;
- Gert Alkmvist (Dpt. of Mathematics, University of Lund, Sweden), who uses the software in his work of enumerating the Calabi–Yau differential equations.

6.2. Multiple binomial sums

Multiple binomial sums form a large class of multi-indexed sequences, closed under partial summation, which contains most of the sequences obtained by multiple summation of binomial coefficients and also all the sequences with algebraic generating function. We study in [14] the representation of the generating functions of binomial sums by integrals of rational functions. The outcome is twofold. Firstly, we show that a univariate sequence is a multiple binomial sum if and only if its generating function is the diagonal of a rational function. Secondly we propose algorithms that decide the equality of multiple binomial sums and that compute recurrence relations for them. In conjunction with geometric simplifications of the integral representations, this approach behaves well in practice. The process avoids the computation of certificates and the problem of accurate summation that afflicts discrete creative telescoping, both in theory and in practice.

6.3. Diagonals of rational functions and selected differential Galois groups

Diagonals of rational functions naturally occur in lattice statistical mechanics and enumerative combinatorics. In all the examples emerging from physics, the minimal linear differential operators annihilating these diagonals of rational functions have been shown to actually possess orthogonal or symplectic differential Galois groups. In order to understand the emergence of such orthogonal or symplectic groups, we exhaustively analyze in [1] three (constrained) sets of diagonals of rational functions, corresponding respectively to rational functions of three variables, four variables and six variables. The conclusion is that, even for these sets of examples which, at first sight, have no relation with physics, their differential Galois groups are always orthogonal or symplectic groups. We also discuss conditions on the rational functions such that the operators annihilating their diagonals do not correspond to orthogonal or symplectic differential Galois groups, but rather to generic special linear groups.

6.4. Algebraic Diagonals and Walks

The diagonal of a multivariate power series F is the univariate power series $\text{Diag}F$ generated by the diagonal terms of F . Diagonals form an important class of power series; they occur frequently in number theory, theoretical physics and enumerative combinatorics. In [7] we study algorithmic questions related to diagonals in the case where F is the Taylor expansion of a bivariate rational function. It is classical that in this case $\text{Diag}F$ is an algebraic function. We propose an algorithm that computes an annihilating polynomial for $\text{Diag}F$. Generically, it is its minimal polynomial and is obtained in time quasi-linear in its size. We show that this minimal polynomial has an exponential size with respect to the degree of the input rational function. We then address the related problem of enumerating directed lattice walks. The insight given by our study leads to a new method for expanding the generating power series of bridges, excursions and meanders. We show that their first N terms can be computed in quasi-linear complexity in N , without first computing a very large polynomial equation. An extended version of this work is presented in [13].

6.5. A human proof of the Gessel conjecture

Counting lattice paths obeying various geometric constraints is a classical topic in combinatorics and probability theory. Many recent works deal with the enumeration of 2-dimensional walks with prescribed steps confined to the positive quadrant. A notoriously difficult case concerns the so-called *Gessel walks*: they are planar walks confined to the positive quarter plane, that move by unit steps in any of the following directions: West, North-East, East and South-West. In 2001, Ira Gessel conjectured a closed-form expression for the number of such walks of a given length starting and ending at the origin. In 2008, Kauers, Koutschan and Zeilberger gave a computer-aided proof of this conjecture. The same year, Bostan and Kauers showed, using again computer algebra tools, that the trivariate generating function of Gessel walks is algebraic. We propose in [3] the first “human proofs” of these results. They are derived from a new expression for the generating function of Gessel walks in terms of special functions. This work has been published in the prestigious journal *Transactions of the AMS*.

6.6. Enumeration of 3-dimensional lattice walks confined to the positive octant

Small step walks in 2D are by now quite well understood, but almost everything remains to be done in higher dimensions. We explored in [2] the classification problem for 3-dimensional walks with unit steps confined to the positive octant. The first difficulty is their number: there are 11 074 225 cases (instead of 79 in dimension 2). In our work, we focused on the 35 548 that have at most six steps. We applied to them a combined approach, first experimental and then rigorous. Among the 35 548 cases, we first found 170 cases with a finite group; in the remaining cases, our experiments suggest that the group is infinite. We then rigorously proved D-finiteness of the generating series in all the 170 cases, with the exception of 19 intriguing step sets for which the nature of the generating function still remains unclear. In two challenging cases, no human proof is currently known, and we derived computer-algebra proofs, thus constituting the first proofs for those two step sets.

6.7. Efficient algorithms for rational first integrals

We presented in [29] fast algorithms for computing rational first integrals with degree bounded by N of a planar polynomial vector field of degree $d \leq N$. The main novelty is that such rational first integrals are obtained by computing via systems of linear equations instead of systems of quadratic equations. This leads to a probabilistic algorithm with arithmetic complexity $\tilde{O}(N^{2\omega})$ and to a deterministic algorithm for solving the problem in $\tilde{O}(d^2 N^{2\omega+1})$ arithmetic operations, where ω is the exponent of linear algebra. By comparison, the best previous algorithm uses at least $d^{\omega+1} N^{4\omega+4}$ arithmetic operations. Our new algorithms are moreover very efficient in practice.

6.8. Quasi-optimal computation of the p -curvature

The p -curvature of a system of linear differential equations in positive characteristic p is a matrix that measures to what extent the system is close to having a fundamental matrix of rational function solutions. This notion, originally introduced in the arithmetic theory of differential equations, has been recently used as an effective tool in computer algebra and in combinatorial applications. We have described in [6] a recent algorithm for computing the p -curvature, whose complexity is almost optimal with respect to the size of the output. The new algorithm performs remarkably well in practice. Its design relies on the existence of a well-suited ring, of so-called Hurwitz series, for which an analogue of the Cauchy–Lipschitz Theorem holds, and on a FFT-like method in which the “evaluation points” are Hurwitz series.

6.9. Axiomatic constraint systems for proof search modulo theories

Goal-directed proof search in first-order logic uses meta-variables to delay the choice of witnesses; substitutions for such variables are produced when closing proof-tree branches, using first-order unification or a theory-specific background reasoner. We have investigated a generalization of such mechanisms whereby theory-specific constraints are produced instead of substitutions. In order to design modular proof-search procedures over such mechanisms, we provide a sequent calculus with meta-variables, which manipulates such constraints abstractly. Proving soundness and completeness of the calculus leads to an acclimatization that identifies the conditions under which abstract constraints can be generated and propagated in the same way unifiers usually are. We then extract from our abstract framework a component interface and a specification for concrete implementations of background reasoners. This is a common work with Damien Rouhling (ENS Lyon), Stéphane Lengrand (CNRS, LIX) and Jean-Marc Notin (CNRS, LIX), based on the PhD contributions of Mahfuza Farooque (unaffiliated). It is described in [8].

6.10. DynaMoW: Dynamic Mathematics on the Web

The interactivity needed by our on-line encyclopedia DDMF is made possible by implementing it over our tool DynaMoW (<http://ddmf.msr-inria.inria.fr/DynaMoW/>). This Ocaml library simultaneously controls external symbolic calculations and web-page generation and was first developed from 2008 to 2011. With the evolution of Ocaml and web technologies, it became possible to hope for a more reactive and configurable tool, by using light-weight threads and websockets. A new design was elaborated this year by F. Chyzak and M. Guesdon, and DynaMoW was rewritten by the latter. Using this new DynaMoW will require a complete and potentially time-consuming port of DDMF. So we decided that experimenting with the port of a smaller DynaMoW-based application should be done to ascertain the new design of DynaMoW-based before going to scale with DDMF. To this end, we applied DynaMoW to another on-line encyclopedia of our's, ECS. The code is now stabilizing, and will be released next year, after documentation is written.

6.11. ECS: Encyclopedia of Combinatorial Structures

The Encyclopedia of Combinatorial Structures (ECS, <http://algo.inria.fr/encyclopedia/>) originates as a project in Project-Team Algorithms, with a first release back in 1998. It is an on-line mathematical encyclopedia with an emphasis on sequences that arise in the context of decomposable combinatorial structures, with the possibility to search by the first terms in the sequence, keyword, generating function, or closed form. As such,

ECS ambitions to be seen as a young cousin of Sloane's famous Encyclopedia of Integer Sequences http://www.research.att.com/articles/featured_stories/2012_03/201203_OEIS.html?fbid=cibE46xiHwx. The latter lists more general types of sequences, and points to numerous entries in ECS for specific properties. With regard to our software development, ECS has served as a nice testbed for several evolutions of DynaMoW, in particular in 2009 and 2011. This year, F. Chyzak and M. Guesdon ported ECS to the language of the new DynaMoW. Public release is expected soon in 2016, and will please the many users waiting for this new release after the former website was discontinued for technical reasons.

6.12. Mathematical Components Library

We have released a new version of the Mathematical Components Library (<http://www.msr-inria.fr/projects/mathematical-components-2/>), including an updated version of the Ssreflect package (<http://ssr.msr-inria.inria.fr/>). A major refactoring of the archive now allows a more modular distribution, through several thematic packages, also available via the OPAM package manager. We have also opened our development repository and we mirror it on the GitHub platform, in order to better foster the community of users of the library.

SUMO Project-Team

7. New Results

7.1. Model expressivity and quantitative verification

7.1.1. *Diagnosability of stochastic systems*

Participants: Nathalie Bertrand, Engel Lefaucheux.

Diagnosis of partially observable stochastic systems prone to faults was introduced in the late nineties. Diagnosability, i.e. the existence of a diagnoser, may be specified in different ways: (1) exact diagnosability (called A-diagnosability) requires that almost surely a fault is detected and that no fault is erroneously claimed while (2) approximate diagnosability (called ϵ -diagnosability) allows a small probability of error when claiming a fault and (3) accurate approximate diagnosability (called AA-diagnosability) requires that this error threshold may be chosen arbitrarily small. In [32] we mainly focus on approximate diagnoses. We first refine the almost sure requirement about finite delay introducing a uniform version and showing that while it does not discriminate between the two versions of exact diagnosability this is no more the case in approximate diagnosis. Then we establish a complete picture for the decidability status of the diagnosability problems: (uniform) ϵ -diagnosability and uniform AA-diagnosability are undecidable while AA-diagnosability is decidable in PTIME, answering a longstanding open question.

7.1.2. *Probabilistic model checking*

Participants: Blaise Genest, Ocan Sankur.

In [16], we considered the verification of Markov chains against properties talking about distributions of probabilities. Even though a Markov chain is a very simple formalism, by discretizing in a finite number of classes the space of distributions through some symbols, we proved that the language of trajectories of distributions (one for each initial distribution) is not regular in general, even with 3 states. We then proposed a parametrized algorithm which approximates what happens to infinity, such that each symbolic block in the approximate language is at most ϵ away from the concrete distribution. We proved in [26] that if the eigenvalues of the Markov chain are distinct positive real numbers, then the trajectory is effectively regular. This is however not the case anymore if the eigenvalues can be distinct roots of real numbers.

Markov decision processes (MDPs) with multi-dimensional weights are useful to analyze systems with multiple objectives that may be conflicting and require the analysis of trade-offs. In [40], we study the complexity of percentile queries in such MDPs and give algorithms to synthesize strategies that enforce such constraints. Given a multi-dimensional weighted MDP and a quantitative payoff function f , thresholds v_i (one per dimension), and probability thresholds α_i , we show how to compute a single strategy to enforce that for all dimensions i , the probability of outcomes ρ satisfying $f_i(\rho) \geq v_i$ is at least α_i . We consider classical quantitative payoffs from the literature (sup, inf, lim sup, lim inf, mean-payoff, truncated sum, discounted sum). Our work extends to the quantitative case the multi-objective model checking problem studied by Etessami et al. [48] in unweighted MDPs.

In the invited contribution [25], we revisit the stochastic shortest path problem, and show how recent results allow one to improve over the classical solutions: we present algorithms to synthesize strategies with multiple guarantees on the distribution of the length of paths reaching a given target, rather than simply minimizing its expected value. The concepts and algorithms that we propose here are applications of more general results that have been obtained recently for Markov decision processes and that are described in a series of recent papers, including [40].

7.1.3. *Stochastic modeling of biological systems*

Participants: Blaise Genest, Éric Fabre, Sucheendra Palaniappan, Matthieu Pichené.

In [47], we model a population of HeLa cells with non deterministic behavior, subject to the drug TRAIL. TRAIL kills a large fraction of cancerous HeLa cells by triggering the apoptosis pathway. Modelling this survival is important to perform *in silico* computations helping designing treatments killing the largest fraction of cancerous cells. We model this system using the stochastic class of Dynamic Bayesian Networks. We maintain large conditional probability tables which are represented by sparse datastructure, and perform simulations by looking ahead one time step and factoring this information to avoid empty probability entries. This considerably improves the simulation based inference of DBNs, getting a 100 times improvement in its efficiency.

7.1.4. Robustness of timed models

Participants: Ocan Sankur, Loïc Hélouët.

Robustness of timed systems aims at studying whether infinitesimal perturbations in clock values can result in new discrete behaviors. A model is robust if the set of discrete behaviors is preserved under arbitrarily small (but positive) perturbations. This year we tackled this problem both for Timed Automata and time Petri Nets.

Timed automata are an extension of finite automata with clock variables that can conveniently model real-time systems. In [42], we study the robustness analysis problem for timed automata under guard imprecisions which consists in computing a timing imprecision bound under which a given specification holds. This is a particular kind of parameter synthesis problems specialized for analyzing robustness. We give a symbolic semi-algorithm for the problem based on a parametric data structure, and evaluate its performance in comparison with a recently published one, and with a binary search on the imprecision bound. We show that a safe bound on imprecision can be computed efficiently, and a performance close to that of exact model checking can be obtained thanks to the use of the parametric data structure and cycle acceleration techniques.

Another related problem is that of robust controller synthesis for timed automata where the goal is to choose actions and their timings so as to ensure a given state is reached when the chosen time delays are adversarially perturbed within a bound. In [21], we are interested in synthesizing “robust” strategies for ensuring reachability of a location in timed automata. We model this problem as a game between the controller and its environment, and solve the parameterized robust reachability problem: we show that the existence of an upper bound on the perturbations under which there is a strategy reaching a target location is EXPTIME-complete. We also extend our algorithm, with the same complexity, to turn-based timed games, where the successor state is entirely determined by the environment in some locations.

We also tackled the robustness problem for time Petri nets (TPNs, for short) in [17] by considering the model of parametric guard enlargement which allows time-intervals constraining the firing of transitions in TPNs to be enlarged by a (positive) parameter. We show that TPNs are not robust in general and checking if they are robust with respect to standard properties (such as boundedness, safety) is undecidable. We then extend the marking class timed automaton construction for TPNs to a parametric setting, and prove that it is compatible with guard enlargements. We apply this result to the (undecidable) class of TPNs which are robustly bounded (i.e., whose finite set of reachable markings remains finite under infinitesimal perturbations): we provide two decidable robustly bounded subclasses, and show that one can effectively build a timed automaton which is timed bisimilar even in presence of perturbations. This allows us to apply existing results for timed automata to these TPNs and show further robustness properties.

7.1.5. Verification for classes of Petri Nets with time

Participants: Blaise Genest, Loïc Hélouët.

We have considered verification problems for classes of Petri Nets with time. We have introduced the first, up to our knowledge, decidability result on reachability and boundedness for Petri Net variants that combine unbounded places, time, and urgency (the ability to enforce actions to happen within some delay). For this, we introduce the class of Timed-Arc Petri Nets with Urgency, which extends Timed-Arc Petri Nets [58] to allow urgency constraints, a feature from Timed-transition Petri Nets (TPNs) [54]. In order to avoid (straightforward) undecidability, we have considered restricted urgency: urgency can be used only on transitions consuming tokens from bounded places. For Timed-Arc Petri Nets with restricted urgency,

we extend decidability results from Timed-Arc Petri Nets: control-state reachability and boundedness are decidable. Our main result concerns (marking) reachability, which is undecidable for both TPNs (because of unrestricted urgency) [52] and Timed-Arc Petri Nets (because of infinite number of clocks) [57]. We have obtained decidability of reachability for (unbounded) TPNs with restricted urgency under a new, yet natural, timed-arc semantics presenting them as Timed-Arc Petri Nets with restricted urgency. Decidability of reachability under the original semantics of TPNs was also obtained for a restricted subclass of unbounded nets. This work is under submission.

7.1.6. Non-interference in partial order models

Participant: Loïc Hélouët.

In [36] we have proposed a new definition of interference for partial order models. Non-interference (NI) is a property of systems stating that confidential actions should not cause effects observable by unauthorized users. Several variants of NI have been studied for many types of models, but rarely for true concurrency or unbounded models. In [36] we have investigated NI for High-level Message Sequence Charts (HMSC), a scenario language for the description of distributed systems, based on composition of partial orders. We firstly have proposed a general definition of security properties in terms of equivalence among observations, and shown that these properties, and in particular NI are undecidable for HMSCs. We hence have considered weaker local properties, describing situations where a system is attacked by a single agent, and show that local NI is decidable in this context. We then have proposed a refinement of local NI to obtain a finer notion of causal NI that emphasizes causal dependencies between confidential actions and observations. This causal NI has then been extended to causal NI with (selective) declassification of confidential events. Finally, we have shown that checking whether a system satisfies local and causal NI and their declassified variants are PSPACE-complete problems. Decidability seems to extend to other classes of partial order models which partially ordered observations can be represented by partial order models that exhibit some forms of regularity such as graph grammars or partial order automata. This conjecture will be explored next year.

7.1.7. Synthesis and games

Participants: Ocan Sankur, Engel Lefauchaux.

In [33], we investigate compositional algorithms to solve safety games described succinctly by synchronous circuits (given by AND and inverter gates). We show how the safety specification can be decomposed, in most cases, into a set of simpler specifications, each defining a safety game depending on less inputs and state variables. We give several algorithms which consist in solving the subgames, and aggregating them in order to find strategies for the global game. We present results of extensive experiments done on around five hundred benchmarks used in the synthesis competition SYNTCOMP 2014 and show that the compositional approach improves the performance on several classes of benchmarks.

In [35] we investigate priced timed games. Priced timed games are two-player zero-sum games played on priced timed automata (whose locations and transitions are labeled by weights modeling the costs of spending time in a state and executing an action, respectively). The goals of the players are to minimise and maximise the cost to reach a target location, respectively. We consider priced timed games with one clock and arbitrary (positive and negative) weights and show that, for an important subclass (the so-called simple priced timed games), one can compute, in exponential time, the optimal values that the players can achieve, with their associated optimal strategies. As side results, we also show that one-clock priced timed games are determined and that we can use our result on simple priced timed games to solve the more general class of so-called reset-acyclic priced timed games (with arbitrary weights and one-clock).

In [34], we introduce a novel rule for synthesis of reactive systems, applicable to systems made of n components which have each their own objectives. This rule is based on the notion of admissible strategies. Intuitively, a strategy σ is dominated by σ' if against all strategies of other players, σ' is as good as σ , and against at least one strategy σ' is strictly better than σ . Admissible strategies are those that are not dominated by any other strategy. The assume-admissible synthesis consists in restricting the space of strategies to admissible ones, and to look for strategy profiles which satisfy given specifications. We compare this rule with previous

rules defined in the literature, and show that contrary to the previous proposals, it defines sets of solutions which are rectangular. This property leads to solutions which are robust and resilient, and allows one to synthesize strategies separately for each agent. We provide algorithms with optimal complexity and also an abstraction framework compatible with the new rule.

7.2. Management of large distributed systems

7.2.1. Parameterized verification in parameterized networks

Participants: Nathalie Bertrand, Paulin Fournier.

We study the problems of reaching a specific control state, or converging to a set of target states, in networks with a parameterized number of identical processes communicating via broadcast. To reflect the distributed aspect of such networks, we restrict our attention to executions in which all the processes must follow the same local strategy that, given their past performed actions and received messages, provides the next action to be performed. We show that the reachability and target problem under such local strategies are NP-complete, assuming that the set of receivers is chosen non-deterministically at each step. On the other hand, these problems become undecidable when the communication topology is a clique. However, decidability can be regained with the additional assumption that all processes are bound to receive the broadcast messages. This is a joint work with Arnaud Sangnier [31].

7.2.2. Runtime enforcement of untimed and timed properties

Participants: Thierry Jéron, Hervé Marchand, Srinivas Pinisetty.

Runtime enforcement is a powerful technique to ensure that a running system satisfies some desired properties. Using an enforcement monitor, an (untrustworthy) input execution (in the form of a sequence of events) is modified into an output sequence that complies with a property. Over the last decade, runtime enforcement has been mainly studied in the context of untimed properties. For several years, and in particular in the context of the PhD thesis of Srinivas Pinisetty [15] we elaborated the theory of runtime enforcement of timed properties. This year we also continued our work on the subject in several directions.

In [38] we describe the TiPEX tool that implements the enforcement monitoring algorithms for timed properties proposed in our previous papers. Enforcement monitors are generated from timed automata specifying timed properties. Such monitors correct input sequences by adding extra delays between events. Moreover, TiPEX also provides modules to generate timed automata from patterns, compose them, and check the class of properties they belong to in order to optimize the monitors. This paper also presents the performance evaluation of TiPEX within some experimental setup.

With colleagues from LaBRI (M. Renard, A. Rollet) and LIG (Y. Falcone) we investigate runtime enforcement of (timed and untimed) properties with uncontrollable events. In [41], we introduce a framework that takes as input any regular (timed) property over an alphabet of events, with some of these events being uncontrollable. An uncontrollable event cannot be delayed nor intercepted by an enforcement mechanism. Enforcement mechanisms satisfy important properties, namely soundness and compliance, meaning that enforcement mechanisms output correct executions that are close to the input execution. We discuss the conditions for a property to be enforceable with uncontrollable events, and we define enforcement mechanisms that modify executions to obtain a correct output, as soon as possible. Moreover, we synthesize sound and compliant descriptions of runtime enforcement mechanisms at two levels of abstraction to facilitate their design and implementation.

With colleagues from the Aalto University (S. Pinisetty, S. Tripakis and V. Preoteasa) and LIG (Y. Falcone) we investigate predictive runtime enforcement. In [39] we introduce predictive runtime enforcement, where the system is not entirely black-box, but we know something about its behavior. This a-priori knowledge about the system allows to output some events immediately, instead of delaying them until more events are observed, or even blocking them permanently. This in turn results in better enforcement policies. We also show that if we have no knowledge about the system, then the proposed enforcement mechanism reduces to a classical non-predictive RE framework. All our results are formalized and proved in the Isabelle theorem prover. We are also currently extending this work to the timed setting.

7.2.3. Discrete controller synthesis

Participants: Nicolas Berthier, Hervé Marchand.

In [29] we investigate the opportunities given by recent developments in the context of Discrete Controller Synthesis algorithms for infinite, logico-numerical systems. To this end, we focus on models employed in previous work for the management of dynamically partially reconfigurable hardware architectures. We extend these models with logico-numerical features to illustrate new modeling possibilities, and carry out some benchmarks to evaluate the feasibility of the approach on such models.

In [30] we elaborate on our former work for the safety control of infinite reactive synchronous systems modeled by arithmetic symbolic transition systems. By using abstract interpretation techniques involving disjunctive polyhedral overapproximations, we provide effective symbolic algorithms allowing to solve the deadlock-free safety control problem while overcoming previous limitations regarding the non-convexity of the set of states violating the invariant to enforce.

The ever growing complexity of software systems has led to the emergence of automated solutions for their management. The software assigned to this work is usually called an Autonomic Management System (AMS). It is ordinarily designed as a composition of several managers, which are pieces of software evaluating the dynamics of the system under management through measurements (e.g., workload, memory usage), taking decisions, and acting upon it so that it stays in a set of acceptable operating states. However, careless combination of managers may lead to inconsistencies in the taken decisions, and classical approaches dealing with these coordination problems often rely on intricate and ad hoc solutions. To tackle this problem, we take a global view and underscore that AMSs are intrinsically reactive, as they react to flows of monitoring data by emitting flows of reconfiguration actions. Therefore in [19] we propose a new approach for the design of AMSs, based on synchronous programming and discrete controller synthesis techniques. They provide us with high-level languages for modeling the system to manage, as well as means for statically guaranteeing the absence of logical coordination problems. Hence, they suit our main contribution, which is to obtain guarantees at design time about the absence of logical inconsistencies in the taken decisions. We detail our approach, illustrate it by designing an AMS for a realistic multi-tier application, and evaluate its practicality with an implementation.

In the invited paper [24] we make an overview of our works addressing discrete control-based design of adaptive and reconfigurable computing systems, also called autonomic computing. They are characterized by their ability to switch between different execution modes w.r.t. application and functionality, mapping and deployment, or execution architecture. The control of such reconfigurations or adaptations is a new application domain for control theory, called feedback computing. We approach the problem with a programming language supported approach, based on synchronous languages and discrete control synthesis. We concretely use this approach in FPGA-based reconfigurable architectures, and in the coordination of administration loops.

7.2.4. Computing knowledge at runtime

Participant: Blaise Genest.

In [37] we compare three notions of knowledge in concurrent system: memoryless knowledge, knowledge of perfect recall, and causal knowledge. Memoryless knowledge is based only on the current state of a process, knowledge of perfect recall can take into account the local history of a process, and causal knowledge depends on the causal past of a process, which comprises the information a process can obtain when all processes exchange the information they have when performing joint transitions. We compare these notions in terms of knowledge strength, number of bits required to store this information, and the complexity of checking if a given process has a given knowledge. We show that all three notions of knowledge can be implemented using finite memory. Causal knowledge proves to be strictly more powerful than knowledge with perfect recall, which in turn proves to be strictly more powerful than memoryless knowledge. We show that keeping track of causal knowledge is cheaper than keeping track of knowledge of perfect recall.

7.2.5. Distributed optimal planning

Participant: Éric Fabre.

Planning problems consist in organizing actions in a system in order to reach one of some target states. The actions consume and produce resources, can of course take place concurrently, and may have costs. We have a collection of results addressing this problem in the setting of distributed systems. This takes the shape of a network of components, each one holding private actions operating over its own resources, and shared/synchronized actions that can only occur in agreement with its neighbors. The goal is to design in a distributed manner a tuple of local plans, one per component, such that their combination forms a consistent global plan of minimal cost.

Our previous solutions to this problem modeled components as weighted automata [22]. In collaboration with Loig Jezequel (TU Munich) and Victor Khomenko (Univ. of Newcastle), we have extended this approach to the case of components modeled as safe Petri nets [23]. This allows one to benefit from the internal concurrency of actions within a component. Benchmarks have shown that this method can lead to significant time reductions to find feasible plans, in good cases. In the least favorable cases, performances are comparable to those obtained with components modeled as automata. The method does not apply to all situations however, as computations require to perform ϵ -reductions on Petri-nets (our work also contains a contribution to this difficult question).

7.2.6. Regulation of urban train systems

Participants: Éric Fabre, Loïc Hérouët, Karim Kecir, Hervé Marchand, Christophe Morvan.

A part of the SUMO team is involved in a collaboration with Alstom transports on regulation techniques. The role of regulation algorithms is to observe train trajectories and delays with respect to an expected ideal schedule, and then compute commands that are sent to trains to meet some quality of service (punctuality, regularity, ...) The objective of this collaboration is to study regulation techniques that are currently in use in urban train systems and compare their performances, and in the future to be able to compute optimal regulation strategies.

This year, we have proposed models inspired from stochastic Petri nets and from closed loop controllers to simulate regulated railways systems. The Petri net model led to the design of a tool called SIMSTORS, that was successfully used to model a real case study (line 1 of Santiago's subway). The simulator relies on event-based symbolic techniques: the time elapsed between two steps of the simulation is the time between two event occurrences (arrival, departure of a train, incident,...). This simulation scheme relying on an abstract model allowed a dramatic speed up of simulation with respect to existing solutions in use at Alstom Transport.

A second line of work has also been explored, in order to design and evaluate new regulation strategies for subway lines. The underlying model is inspired from event-based control theory, in a stochastic and timed setting. It abstracts away several significant topological features of a subway line, and focuses on the optimal command of train speeds in order to achieve high-level objectives such as the equal spacing of trains, or the efficient insertion/extraction of trains. This approach has allowed us to design new distributed regulation policies, which are remarkably stable and efficiently mitigate known instabilities of subway lines, like the bunching phenomenon. We are currently working on an extension of this approach for the management of time-tables and of forks and joins in the topology of subway lines.

7.3. Data driven systems

7.3.1. A model of large-scale distributed collaborative system

Participants: Éric Badouel, Loïc Hérouët, Christophe Morvan, Robert Nsaibirni.

We have presented in [27] and [18] a purely declarative approach to artifact-centric collaborative systems, a model which we introduced in two stages. First, we assume that the workspace of a user is given by a mindmap, shortened to a map, which is a tree used to visualize and organize tasks in which he or she is involved, together with the information used for the resolution of these tasks. We introduce a model of guarded attribute grammar, or GAG, to help the automation of updating such a map. A GAG consists of an underlying grammar, that specifies the logical structure of the map, with semantic rules which are used both to govern the evolution of the tree structure (how an open node may be refined to a subtree) and to compute the value of some of the attributes (which derives from con-textual information). The map enriched with this extra information

is termed an active workspace. Second, we define collaborative systems by making the various user's active workspaces communicate with each other. The communication uses message passing without shared memory thus enabling convenient distribution on an asynchronous architecture. A case study on a disease surveillance system is under development in the PhD thesis of Robert Nsaibirni and a first prototype of the model of active workspaces was written by Eric Badouel.

7.3.2. *Petri Nets with semi-structured data*

Participants: Éric Badouel, Loïc Hélouët, Christophe Morvan.

In [28], we have proposed an extension of Petri nets with data called Structured Data Nets (StDN). This extension allows for the description of transactional systems with data. In StDNs, tokens are structured documents. Each transition is attached to a query, guarded by patterns, (logical assertions on the contents of its preset) and transforms tokens. In [28], we have proposed a semantics for StDNs, and then considered their formal properties: coverability of a marking, termination and soundness of transactions. Unrestricted StDNs are Turing complete, so these properties are undecidable. However, we have proposed an order on structured documents, and shown that under reasonable restrictions on documents and on the expressiveness of patterns and queries, StDNs are well-structured transition systems, for which coverability, termination and soundness are decidable. This work has then been extended to consider properties of sets of configurations described as upward closed sets satisfying patterns, and should appear in a journal paper in 2016.

TASC Project-Team

7. New Results

7.1. IBEX

The development of the Ibox library has continued. The main developments in 2015 are:

- the complete refactoring of the multi-heap internal structure used for search space exploration in the global optimizer
- the creation of a new module for explicit set (or pavings) manipulation/algebra with full documentation and tutorial

7.2. NetWMS2

New advances have been made in the context of packing curved objects. The packing algorithm developed in 2014 have been published in ICJAI'15, along with new features. The calculation of the *penetration depth* (a classical measure of violation cost for overlapping objects) has also been extended to the case of parametric curves (like, e.g., Bezier curves) and new experiments have been conducted with our solver for this new type of objects.

We deal with the problem of packing two-dimensional objects of quite arbitrary shapes including in particular curved shapes (like ellipses) and assemblies of them. This problem arises in industry for the packaging and transport of bulky objects which are not individually packed into boxes, like car spare parts. There has been considerable work on packing curved objects but, most of the time, with specific shapes; one famous example being the circle packing problem. There is much less algorithm for the general case where different shapes can be mixed together. A successful approach has been proposed recently by Martinez et al. and the algorithm we propose here is an extension of their work. Martinez et al. use a stochastic optimization algorithm with a fitness function that gives a violation cost and equals zero when objects are all packed. Their main idea is to define this function as a sum of $\binom{n}{2}$ elementary functions that measure the overlapping between each pair of different objects. However, these functions are ad-hoc formulas. Designing ad-hoc formulas for every possible combination of object shapes can be a very tedious task, which dramatically limits the applicability of their approach. We generalize the approach by replacing the ad-hoc formulas with a numerical algorithm that automatically measures the overlapping between two objects. Then, we come up with a fully black-box packing algorithm that accept any kind of objects.

7.3. Time-Series Constraints

We describe a large family of constraints for structural time series by means of function composition. These constraints are on aggregations of features of patterns that occur in a time series, such as the number of its peaks, or the range of its steepest ascent. The patterns and features are usually linked to physical properties of the time series generator, which are important to capture in a constraint model of the system, i.e. a conjunction of constraints that produces similar time series. We formalise the patterns using finite transducers, whose output alphabet corresponds to semantic values that precisely describe the steps for identifying the occurrences of a pattern. Based on that description, we automatically synthesise automata with accumulators, as well as constraint checkers. The description scheme not only unifies the structure of the existing 30 time-series constraints in the Global Constraint Catalogue, but also leads to over 600 new constraints, with more than 100,000 lines of synthesised code.

7.4. New Global Constraints

This year we introduce new generic global constraints that can be respectively used to reformulate a number of constraints where the formulation become easy once some tuples are sorted, and to express temporal relation between two sequence of intervals.

- Some constraint programming solvers and constraint modelling languages feature the $sort(L, P, S)$ constraint, which holds if S is a nondecreasing rearrangement of the list L , the permutation being made explicit by the optional list P . However, such sortedness constraints do not seem to be used much in practice. We argue that reasons for this neglect are that it is impossible to require the underlying sort to be stable, so that $sort$ cannot be guaranteed to be a total-function constraint, and that L cannot contain tuples of variables, some of which form the key for the sort. To overcome these limitations, we introduce the *stable-key-sort* constraint, decompose it using existing constraints, and propose a propagator. This new constraint enables a powerful modelling idiom, which we illustrate by elegant and scalable models of two problems that are otherwise hard to encode as constraint programs.
- The constraint was initially motivated by an application where the objective is to generate a video summary, built using intervals extracted from a video source. In this application, the constraints used to select the relevant pieces of intervals are based on Allen's algebra. The best state-of-the-art results are obtained with a small set of ad hoc solution techniques, each specific to one combination of the 13 Allen's relations. Such techniques require some expertise in Constraint Programming. This is a critical issue for video specialists. We design a generic constraint, dedicated to a class of temporal problems that covers this case study, among others. *ExistAllen* takes as arguments a vector of tasks, a set of disjoint intervals and any of the 213 combinations of Allen's relations. *ExistAllen* holds if and only if the tasks are ordered according to their indexes and for any task at least one relation is satisfied, between the task and at least one interval. We design a propagator that achieves bound-consistency in $O(n+m)$, where n is the number of tasks and m the number of intervals. This propagator is suited to any combination of Allen's relations, without any specific tuning. Therefore, using our framework does not require a strong expertise in Constraint Programming. The experiments, performed on real data, confirm the relevance of our approach.

7.5. Controlling the Generation of Solutions

The following two results deal with controlling the generation of solutions to a constraint problem.

- The *focus* constraint expresses the notion that solutions are concentrated. In practice, this constraint suffers from the rigidity of its semantics. To tackle this issue, we propose three generalizations of the FOCUS constraint. We provide for each one a complete filtering algorithm. Moreover, we propose mathematical programming (ILP) and constraint programming decompositions.
- There are significant motivations for considering alternate solutions to a problem. As expressed by renowned statistician George Box *The most that can be expected from any model is that it can supply a useful approximation to reality: all models are wrong; some models are useful.* Multiple solutions alone, however, are not sufficient to guarantee anything of value. If they are nearly identical nothing is gained. While most frameworks in the literature consider diversity between solutions through mathematical distances, this paper proposes alternative distance measures represented by global constraints. It introduces a constraint programming framework for optimization problems, able to generate sets of nearly-optimal solutions that are diverse. With respect to over-constrained problems, the framework can be specialized in order to generate solution sets where constraint violations are diverse.

TEA Project-Team

7. New Results

7.1. Polychronous automata

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

We have defined a model of *polychronous automata* based on clock relations [13]. A specificity of this model is that an automaton is submitted to clock constraints: these finite-state automata define transition systems to express explicit reactions together with properties, in the form of Boolean formulas over logical time, to constrain their behavior. This allows one to specify a wide range of control-related configurations, either reactive, or restrictive with respect to their control environment. A semantic model is defined for these polychronous automata, that relies on a Boolean algebra of clocks. Polychronous automata integrate smoothly with data-flow equations in the polychronous model of computation.

This formal model of automata also supports the recommendations adopted by the SAE committee on the AADL to implement a timed and synchronous behavioural annex for the standard⁰.

A minimal syntactic extension of the Signal language has been defined to integrate polychronous automata in Polychrony. We have added a new syntactic category of *process*, called *automaton*. In such an automaton process, labeled processes represent states, and generic processes such as `Transition` are used to represent the automaton features. Usual equations can be used in these automaton processes to specify constraints or to define computations.

We have also defined and implemented the refinement of Signal processes as automata. A given Signal program may be seen as an automaton which contains one single state and one single transition, labeled by a clock. This clock is the upper bound of all the clocks of the program (the *tick* of the program). The construction of a refined automaton from a Signal program is based on delayed signals, viewed as state variables (in particular Boolean ones). A state of the automaton is a Signal program with some valuation of its state variables. Transitions are labeled by clocks, which represent the events that fire these transitions. The principle of the construction consists in dividing a given state according to the possible values of a state variable (i.e., *true* and *false* for Boolean state variables) in order to get two states, and thus two new Signal programs. Each one of these two states is obtained using a rewriting of the starting program. Moreover, the absence of value for the state variable (which can be considered as another possible value) is taken into account in the clocks labelling the transitions. The construction of the automaton is a hierarchic process. Thanks to the clock hierarchy, this construction, which would be expensive in the worst case (the size of the explicit automaton being an exponential of its number of state variables), may be heavily simplified.

7.2. Runtime verification and trace analysis

Participants: Vania Joloboff, Daian Yue, Frédéric Mallet.

When engineers design a new cyber physical system, there are well known requirements that can be translated as system properties that must be verified. These properties can be expressed in some formalism and when the model has been designed, the properties can be checked at the model level, using model checking techniques or other model verification techniques. When building a virtual prototype of the system, including a combination of simulated hardware, firmware and application software, the executable models can be augmented also with property verification, for example in the PSL language, or simply by introducing assertions in the implementation code.

⁰Logically timed specification in the AADL: a synchronous model of computation and communication (recommendations to the SAE committee on AADL). L. Besnard, E. Borde, P. Dissaux, T. Gautier, P. Le Guernic, and J.-P. Talpin. Technical Report RT-0446, Inria, 2014.

This requires that the properties are well specified at the time the virtual prototype is assembled. However it is also the case that many intrinsic properties are actually unforeseen when the virtual prototype is assembled, for example that some hardware buffer overflow should not remain unnoticed by the software. In most cases, during system design the simulation fails: the engineers then must investigate the cause of the failure. Most of time the failure is due to an unexpected sequence of states and transitions that involve several components mixing hardware and software that could not be checked at the model level (e.g. state explosion) or was simply unforeseen. The engineers then have to investigate the cause of failure.

A widely used technique for that consists in storing all of the trace data of simulation sessions into trace files, which are analyzed later with specialized trace analyzer tools. Such trace files have become huge, possibly hundred of Gigabytes as all data are stored into the trace files, and have become untractable by human manual handling. The engineers use some kind of search tools to identify the cause of failure and after iterative refinement steps, which are very time consuming, eventually identify the reason, most often some unforeseen causality chain of events and state transitions that lead to a failure. A new system property can then be captured and included into the set of verified properties.

In order to better identify the reason for such failures and capture the missing properties that the system should verify we have started to work on a new run time verification approach based on trace analysis. Approaches like PSL requires that the properties to verify are known before hand. Our approach is attempting for the engineers to experiment various property verification of failing simulations without re-building the virtual prototype. We are investigating a technique for trace analysis that makes it possible to investigate properties either statically working from a trace file or dynamically by introducing a dynamic verification component into the virtual prototype.

The first idea is to introduce a formal mapping/filtering technique such that the raw data generated by a virtual prototype can be mapped onto a formal trace model. For that, we propose to use a model transformer whose code is generated from a higher level. Using the Eclipse modelling framework, we propose for the virtual prototyping engineers to first describe using a Domain Specific Language how the raw output of the simulator can be filtered and mapped to a formal model. This Domain Specific Language takes as input the description of the simulator output, and the description of the formal output, following fixed meta models. In current version, the meta model of the virtual prototype dictates that it generates 'trace items' where each trace item is specified as a sequence of identified binary data variables (bits, bytes, words..) that carry a timestamp.

The model transformer generates code (in our case C++) that is dynamically invoked by the virtual prototype to dynamical map the trace output. An advantage of doing that is that all irrelevant data with regards to a tested property can be ignored and the size of trace files can be considerably reduced. For our experiment, we have chosen logical clock CCSL as our formal target formalism. The Eclipse EMF tool we have defined allows users to define a mapping model from the local simulation events from the SimSoC simulator to a logical clock format.

The second idea is to hide the complexity of the formal method formulas into a user friendly property specification language. For example, we do not want to expose the end-users engineers to understand the intricacies of CCSL or LTL. The property specification language is translated into CCSL formulas, which in turn generate automata. It should be possible then, to some extent, to change the formalism underneath the language without changing the properties expressed by the user.

The property specification language ultimately compiles into automata that parse the formal trace output generated above. At runtime of the virtual prototype, the mapping library is dynamically loaded by the simulator and generates input for the automata. The verification of the properties can be dynamic, with a true runtime verification, or statically by analyzing the (much smaller) trace file after a failure.

This year we have investigated this approach, designed the architecture described above and carried some experimental work, but a significant part of the implementation still remains to be done. We have started designing a new property specification language where the users can express properties such as causality (e.g. the train must not start if the door is opened) or jittering or clock drift in image processing [11], [10]. There remain some theoretical issues with regards to which properties can be effectively verified.

7.3. Integration of Polychrony with QGen

Participants: Christophe Junke, Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

The FUI project P gave birth to the QGen qualifiable model compiler, developed by Adacore. The tool accepts a discrete subset of Simulink expressed in a language called P and produces C or Ada code. It is currently not known if an architectural description language is going to be integrated in QGen, as originally planned.

We developed a transformation tool named P2S for expressing P system models in Signal, using the EMF (Eclipse Modelling Framework) technology. P2S tool is written in Clojure, a language inspired by Lisp running on the Java Virtual Machine, which helped us define a terse and expressive API for manipulating Signal models while remaining fully interoperable with existing Java libraries (including Eclipse plugins and especially Polychrony ones).

We experimented this transformation tool on small to medium use cases provided by members of the P project. Our work is detailed in a conference paper titled “Integration of Polychrony and QGen Model Compiler”, which will appear at ERTSS’16⁰. A perspective of our work is to convert the intermediate code emitted by QGen as Signal too (under development), in order to produce a fully executable Signal model of Simulink models, and combine them with architectural description of systems in AADL, and/or P’s architecture language.

7.4. Formal semantics and model-based analysis of AADL specifications

Participants: Loïc Besnard, Etienne Borde, Thierry Gautier, Paul Le Guernic, Clément Guy, Jean-Pierre Talpin, Huafeng Yu.

Last year, the SAE committee on the AADL adopted our recommendations to implement a timed and synchronous behavioural annex for the standard. We have defined a new model of polychronous constrained automata that has been provided as semantic model for our proposal of an extension of the AADL behavioural annex. An experimental implementation of the semantic features of this “timing annex” will be provided through the Polychrony framework. For that purpose, representations of automata have been introduced in the Signal toolbox of Polychrony. The implementation will enrich the already existing transformation from AADL models to Signal programs to consider behaviour of AADL models, and will be integrated in the POP environment for Eclipse. The transformation from AADL behaviour annex to Signal programs use the Signal extension for polychronous automata, which are used as the common semantic domain. The implementation is currently tested with the adaptive cruise control case study developed with Toyota ITC.

Our work with the SAE committee is sponsored by Toyota, with whom we started a new project in 2014 jointly with VTRL as US partner. The main topic of our project is the semantic-based model integration of automotive architectures, virtual integration, toward formal verification and automated code synthesis [19]. The project led to the elaboration of a case study of an adaptive cruise control system, supported through an AADL implementation and a video of demonstration. The case study implementation is an AADL model representing the whole adaptive cruise control system, from car devices (e.g., brakes, throttle or radar) to software behavior, including embedded hardware (buses, processors and memories). It will be used in the future to demonstrate property and constraint analyses through heterogeneous systems. Huafeng Yu, our main collaborator at Toyota ITC, presented the video of demonstration at the annual Toyota show case. Early returns from the show case express a growing interest of Toyota for architecture and timing of car embedded systems, which could lead to new collaborations.

7.5. Refinement types for reactive system models

Participants: Pierre Jouvelot, Sandeep Shukla, Jean-Pierre Talpin.

⁰Integration of polychrony in the QGen model compiler. C. Junke, T. Gautier, L. Besnard, J.-P. Talpin. ERTS’16 - European Congress on Embedded Real-Time Software and Systems, 2016.

We introduced a new technique born from the field of functional programming to adapt and extend it to the case of reaction systems, the notion of refinement types of Jahla et al.⁰. Our idea is to formulate the analysis of algebraic properties in synchronous and reactive programs as data-dependent type properties formulated using multi-sorted logic formulas, which we call liquid clocks [20], [18]. Our objectives are to cover the case of several models of concurrency and computation: synchronous, asynchronous, data-parallel; as well as to formulate such algebraic properties for linear, continuous and logical forms of time, all into the same type-theoretical framework. This work, born from two collaborations With USAF/VT and with the ANR Feever project, will be pursued within the TIX international partnership.

7.6. Formal verification of timing aspects of cyber-physical systems using a contract theory

Participants: Jean-Pierre Talpin, Benoit Boyer, David Mentre, Simon Lunel.

This is a new project in collaboration with Mitsubishi Electronics Research Centre Europe (MERCE). The primary goal of our project is to ensure correctness-by-design in cyber-physical systems, i.e., systems that mix software and hardware in a physical environment, e.g., Mitsubishi factory automation lines. We plan to explore a multi-sorted algebraic framework for static analysis and formal verification starting from a simple use case extracted from Mitsubishi factory automation documentations. This will serve as a basis to more ambitious research where we intend to leverage recent advance in type theory, SMT solvers for nonlinear real arithmetic (dReal and δ -decidability) and contracts theory (meta-theory of Benveniste et al., Ruchkin's contracts) to provide a general framework of reasoning about heterogeneous factory components.

⁰*Liquid Types*. P. M. Rondon, M. Kawaguchi, R. Jhala. PLDI, 2008

TOCCATA Project-Team

7. New Results

7.1. Deductive Verification

- M. Clochard, J.-C. Filliâtre, and A. Paskevich proposed a novel method to prove the relative safety of operations over bounded integers in a large class of programs. Their approach consists of introducing dedicated abstract types for the bounded integers and restricting the set of allowed operations over these types in such a way that it is impossible to reach the bound during a realistic execution of the program: for example, it would take several hundred years to overflow a 64-bit integer. This technique is aimed at integer variables that serve essentially as counters or size measures. It can be used alongside the traditional methods of proving the absence of overflows for other integer values in the same program. The proposed approach is implemented in Why3 and was presented at VSTTE 2015 [26].
- J.-C. Filliâtre and M. Pereira proposed a new way to specify the behavior of a cursor data structure, with the objective of being able to verify both the implementation of a cursor and its use by client code. The approach is modular, which means that a program using a cursor can be verified independently of the way the cursor is implemented. An experimental evaluation has been conducted with Why3, with several implementations and client codes being verified. This work will be presented at JFLA 2016 [26].
- C. Fumex and C. Marché developed a new library for bit-vectors, in Why3 and SPARK [30]. This library is rich enough for the formal specification of functional behavior of programs that operate at the level of bits. It is also designed to exploit efficiently the support for bit-vectors built-in in some SMT solvers. This work is done in the context of the ProofInUse joint laboratory. The SPARK front-end of Why3, for the verification of Ada programs, is extended to exploit this new bit-vector theory. Several cases studies are conducted: efficient search for rightmost bit of a bit-vector, efficient computation of the number of bits set to 1, efficient solving of the n -queens problem. At the level of SPARK, a program inspired from some industrial code (originally developed in C par J. Gerlach, Fraunhofer FOKUS Institute, Germany and partially proved with Frama-C and Coq) is specified in SPARK and proved with automatic solvers only. The support for bit-vectors is already distributed with SPARK, and SPARK users already reported that several verification conditions, that couldn't be proved earlier, are now proved automatically.
- D. Hauzar and C. Marché worked on counterexample generation from failed proof attempts. They designed a new approach for generating potential counterexamples in the deductive verification setting, and implemented in Why3. When the logic goal generated for a given verification condition is not shown unsatisfiable by an SMT solvers, some solver can propose a model. By carefully reverting the transformation chain (from an input program through the VC generator and the various translation steps to solvers), this model is turned into a potential counterexample that the user can exploit to analyze why its original code is not proved. The approach is implemented in the chain from Ada programs through SPARK, Why3, and SMT solvers CVC4 and Z3. This implementation is robust enough to be distributed in the next release Pro 16 of SPARK. A research report on this subject will appear in January 2016.
- A. Charguéraud and F. Pottier (Inria Paris-Rocquencourt) obtained new results in the machine-checked verification of asymptotic complexity bounds, in addition to program correctness properties. Verifying the time usage of a program is very important, because otherwise a program might be proved to be functionally correct but may appear to run into an infinite loop for particular input data. More specifically, A. Charguéraud and F. Pottier started from the extension of CFML with *time credits* (encoding of time resources in Separation Logic), developed last year by A. Charguéraud, and

they used it to formally produce a machine-checked proof of the correctness and time complexity of a Union-Find data structure, implemented as an OCaml module. They thereby demonstrate that the approach scales up to difficult complexity analyses, and applies to actual executable code (as opposed to pseudo-code). This work was presented at ITP 2015 [24]. Furthermore, A. Charguéraud and F. Pottier co-advised the M2 internship of Armaël Guéneau, who extended the time credits approach so as to allow working conveniently with the big- O notation. He extended the CFML library and verified the time complexity of a binary random access list data structure due to Okasaki. This work has not been published yet.

- A. Charguéraud described a method for reasoning about mutable data structures that own their elements. In Separation Logic, representation predicates describe the ownership of a mutable data structure, by establishing a relationship between the entry point of the structure, the piece of heap over which this structure spans, and the logical model associated with the structure. When a data structure is polymorphic, such as in the case of a container, its representation predicate needs to be parameterized not just by the type of the items stored in the structure, but also by the representation predicates associated with these items. Such higher-order representation predicates can be used in particular to control whether containers should own their items. A. Charguéraud wrote a paper describing, through a collection of practical examples, solutions to the challenges associated with reasoning about accesses into data structures that own their elements. This paper will appear at CPP 2016 [23].

7.2. Automated Reasoning

- C. Dross, A. Paskevich, J. Kanig and S. Conchon published a journal paper [16] about integration of first-order axiomatizations with triggers as decision procedures in an SMT solver. This work extends a part of C. Dross PhD thesis [79]. A formal semantics of the notion of trigger is presented, with a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating. An extended DPLL(T) algorithm can then integrate such an axiomatization with triggers, as a decision procedure for the theory it defines.

7.3. Certification of Languages, Tools and Systems

- M. Clochard and L. Gondelman developed a formalization of a simple compiler in *Why3*. It compiles a simple imperative language into assembler instructions for a stack machine. This case study was inspired by a similar example developed using Coq and interactive theorem proving. The aim is to improve significantly the degree of automation in the proofs. This is achieved by the formalization of a Hoare logic and a Weakest Precondition Calculus on assembly programs, so that the correctness of compilation is seen as a formal specification of the assembly instructions generated. This work was presented at the JFLA conference in 2015 [25].
- S. Boldo, C. Lelay, and G. Melquiond worked on the Coquelicot library, designed to be a user-friendly Coq library about real analysis. An easier way of writing formulas and theorem statements is achieved by relying on total functions in place of dependent types for limits, derivatives, integrals, power series, and so on. To help with the proof process, the library comes with a comprehensive set of theorems and some automation. We have exercised the library on several use cases: in an exam at university entry level, for the definitions and properties of Bessel functions, and for the solution of the one-dimensional wave equation. These results are published in the journal *Mathematics in Computer Science* [14].
- C. Lelay developed a new formalization of convergence with a focus on usability and genericity for the Coquelicot library. This formalization covers various parts of analysis: sequences, real functions, complex functions, vector functions, and so on. This work was presented at the 7th Coq Workshop [27].
- C. Paulin wrote a gentle introduction to the Calculus of Inductive Construction, the formalism on which the Coq proof assistant is based [28], discussing both theoretical and pragmatic aspects of the design.

7.4. Floating-Point and Numerical Programs

- É. Martin-Dorel and G. Melquiond worked on integrating the CoqInterval and CoqApprox libraries into a single package. The CoqApprox library is dedicated to computing verified Taylor models of univariate functions so as to compute approximation errors. The CoqInterval library reuses this work to automatically prove bounds on real-valued expressions. A large formalization effort took place during this work, so as to get rid of all the holes remaining in the formal proofs of CoqInterval. It was also the chance to perform a comparison between numerous decision procedures dedicated to proving nonlinear inequalities involving elementary functions. This work has been published in the *Journal of Automated Reasoning* [18].
- S. Boldo and G. Melquiond, with J.-H. Jourdan and X. Leroy (Gallium team, Inria Paris - Rocquencourt) extended the CompCert compiler to get the first formally verified C compiler that provably preserves the semantics of floating-point programs. This work, published in the *Journal of Automated Reasoning* [13], also covers the formalization of numerous algorithms of conversion between integers and floating-point numbers.
- S. Boldo worked on the fact that $a/\sqrt{a^2 + b^2}$ is always in the interval $[-1, 1]$ even when operations are done using floating-point arithmetic. This reduces to taking the square root of the square of a floating-point number as it is the worst case. Results in radix 2 (where $\sqrt{a^2} = |a|$) and other radices (where it might not hold) have been published at the 8th International Workshop on Numerical Software Verification [22].
- S. Boldo worked on programs computing the average of two floating-point numbers. As we want to take exceptional behaviors into account, we cannot use the naive formula $(x+y)/2$. Based on hints given by Sterbenz, she first wrote an accurate program and formally proved its properties. She also developed and formally proved a new algorithm that computes the correct rounding of the average of two floating-point numbers [21]. This was published at the 17th International Conference on Formal Engineering Methods.
- P. Roux formalized a theory of numerical analysis for bounding the round-off errors of a floating-point algorithm. This approach was applied to the formal verification of a program for checking that a matrix is semi-definite positive. The challenge here is that testing semi-definiteness involves algebraic number computations, yet it needs to be implemented using only approximate floating-point operations. This work has been published in the *Journal of Automated Reasoning* [19].
- C. Lelay and G. Melquiond worked on formalizing in Coq a numerical domain for the Verasco abstract interpreter built upon the CompCert verified compiler. This abstract domain is a relational domain based on affine forms (zonotopes). It is meant to help verifying floating-point programs and it is expected to perform faster (but less accurately) than a more generic domain based on polyhedrons.

7.5. Miscellaneous

- A. Charguéraud worked together with Umut Acar, Mike Rainey, and Filip Sieczkowski, as part of the ERC project *DeepSea*, on the development of efficient data structures and algorithms targeting modern, shared memory multicore architectures. A. Charguéraud was involved in two major results obtained this year.

The first result is the development of fast and robust parallel graph traversal algorithms based on depth-first-search. This algorithm leverages a new sequence data structure for representing the set of edges remaining to be visited. This sequence itself builds on prior work on bootstrapped chunked sequences [35]. In particular, the edge sequence structure uses a balanced split operation for partitioning the edges of a graph among the several processors involved in the computation. Compared with prior work, the new algorithm is designed and proved to be efficient not just for particular classes of graphs, but for all input graphs. This work has been published in the ACM/IEEE Conference on High Performance Computing (SC) [20].

Another result by A. Charguéraud and his co-authors is the development of a calculus for parallel computing on shared memory computers. Many languages for writing parallel programs have been developed. These languages offer several different abstractions for parallelism, such as fork-join, async-finish, futures, etc. While they may seem similar, these abstractions lead to different semantics, language design and implementation decisions. In this work, we consider the question of whether it would be possible to unify different approaches to parallelism. To this end, we propose a calculus, called *DAG-calculus* that can encode existing approaches to parallelism based on fork-join, async-finish, and futures paradigms and possibly others. We have shown that the approach is realistic by presenting an implementation in C++ and by performing an empirical evaluation. This work has been submitted for publication.

- A. Charguéraud developed a patch to the OCaml compiler for improving type error messages, in particular to make the language more accessible to beginners. The problem of improving type error messages in ML has received quite a bit of attention over the past two decades, and many different strategies have been considered. The challenge is not only to produce error messages that are both sufficiently concise and systematically useful to the programmer, but also to handle a full-blown programming language and to cope with large-sized programs efficiently. A. Charguéraud's novel approach consists of a slight modification to the traditional ML type inference algorithm implemented in OCaml that, by significantly reducing the left-to-right bias, produces error messages that are more helpful to the programmer. This work was published this year in the journal *Electronic Proceedings in Theoretical Computer Science* [15].

VEGAS Project-Team

6. New Results

6.1. Robustness issues in computational geometry

Participants: Olivier Devillers, Monique Teillaud.

6.1.1. Qualitative Symbolic Perturbation: a new geometry-based perturbation framework

In a classical Symbolic Perturbation scheme, degeneracies are handled by substituting some polynomials in ϵ to the input of a predicate. Instead of a single perturbation, we propose to use a sequence of (simpler) perturbations. Moreover, we look at their effects geometrically instead of algebraically; this allows us to tackle cases that were not tractable with the classical algebraic approach [25].

This work was done in collaboration with Menelaos Karavelas (Univ. of Crete).

6.2. Probabilistic analysis of geometric data structures and algorithms

Participant: Olivier Devillers.

6.2.1. The worst visibility walk in a random Delaunay triangulation is $O(\sqrt{n})$

We show that the memoryless routing algorithms Greedy Walk, Compass Walk, and all variants of visibility walk based on orientation predicates are asymptotically optimal in the average case on the Delaunay triangulation. More specifically, we consider the Delaunay triangulation of an unbounded Poisson point process of unit rate and demonstrate that the worst-case path between any two vertices inside a domain of area n has a number of steps that is not asymptotically more than the shortest path which exists between those two vertices with probability converging to one (as long as the vertices are sufficiently far apart.) As a corollary, it follows that the worst-case path has $O(\sqrt{n})$ steps in the limiting case, under the same conditions. Our results have applications in routing in mobile networks and also settle a long-standing conjecture in point location using walking algorithms. Our proofs use techniques from percolation theory and stochastic geometry [24].

This work was done in collaboration with Ross Hemsley (formerly in Inria Geometrica).

6.2.2. Smooth analysis of convex hulls

We establish an upper bound on the smoothed complexity of convex hulls in \mathbb{R}^d under uniform Euclidean (ℓ^2) noise. Specifically, let $\{p_1^*, p_2^*, \dots, p_n^*\}$ be an arbitrary set of n points in the unit ball in \mathbb{R}^d and let $p_i = p_i^* + x_i$, where x_1, x_2, \dots, x_n are chosen independently from the unit ball of radius δ . We show that the expected complexity, measured as the number of faces of all dimensions, of the convex hull of $\{p_1, p_2, \dots, p_n\}$ is $O\left(n^{2-\frac{4}{d+1}}(1+1/\delta)^{d-1}\right)$; the magnitude δ of the noise may vary with n . For $d = 2$ this bound improves to $O\left(n^{\frac{2}{3}}(1+\delta^{-\frac{2}{3}})\right)$.

We also analyze the expected complexity of the convex hull of ℓ^2 and Gaussian perturbations of a nice sample of a sphere, giving a lower-bound for the smoothed complexity. We identify the different regimes in terms of the scale, as a function of n , and show that as the magnitude of the noise increases, that complexity varies monotonically for Gaussian noise but non-monotonically for ℓ^2 noise [13].

This work was done in collaboration with Xavier Goaoc (Univ. Marne la Vallée), Marc Glisse and Remy Thomasse (Inria Geometrica).

6.3. Non-linear computational geometry

Participants: Guillaume Moroz, Sylvain Lazard, Marc Pouget, Laurent Dupont, Rémi Imbach.

6.3.1. Solving bivariate systems and topology of plane algebraic curves

In the context of our algorithm Isotop for computing the topology of plane algebraic curves (see Section 5.1), we work on the problem of solving a system of two bivariate polynomials. We are interested in certified numerical approximations or, more precisely, isolating boxes of the solutions. But we are also interested in computing, as intermediate symbolic objects, a Rational Univariate Representation (RUR) that is, roughly speaking, a univariate polynomial and two rational functions that map the roots of the univariate polynomial to the two coordinates of the solutions of the system. RURs are relevant symbolic objects because they allow to turn many queries on the system into queries on univariate polynomials. However, such representations require the computation of a separating form for the system, that is a linear combination of the variables that takes different values when evaluated at the distinct solutions of the system.

We published this year [11] results showing that, given two polynomials of degree at most d with integer coefficients of bitsize at most τ , (i) a separating form, (ii) the associated RUR, and (iii) isolating boxes of the solutions can be computed in, respectively, $\tilde{O}_B(d^8 + d^7\tau)$, $\tilde{O}_B(d^7 + d^6\tau)$ and $\tilde{O}_B(d^8 + d^7\tau)$ bit operations in the worst case, where \tilde{O} refers to the complexity where polylogarithmic factors are omitted and O_B refers to the bit complexity.

However, during the publishing process, we have substantially improved these results. We have presented for these three sub-problems new algorithms that have worst-case bit complexity $\tilde{O}_B(d^6 + d^5\tau)$. We have also presented probabilistic Las Vegas variants of our two first algorithms, which have expected bit complexity $\tilde{O}_B(d^5 + d^4\tau)$. We also show that it is likely difficult to improve these complexities as it would essentially require to improve bounds on other fundamental problems (e.g., computing resultants, checking squarefreeness and root isolation of univariate polynomials) that have hold for decades.

This work was done in collaboration with Yacine Bouzidi (Inria Saclay), Michael Sagraloff (MPII Sarrebrücken, Germany) and Fabrice Rouillier (Inria Rocquencourt). It is published in the research report [22] and submitted to a journal.

A key ingredient of the above work is the classical triangular decomposition algorithm via subresultants [31] on which we obtain two results of independent interest. First, we improved by a factor d the state-of-the-art worst-case bit complexity of this algorithm [22]. One constraint on this algorithm is that it requires that the curves defined by the input polynomials have no common vertical asymptotes. Our second result is a generalization of this algorithm, which removes that restriction while preserving the same worst-case bit complexity of $\tilde{O}_B(d^6 + d^5\tau)$. Furthermore, we actually present a refined bit complexity in $\tilde{O}_B(d_x^3 d_y^3 + (d_x^2 d_y^3 + d_x d_y^4)\tau)$ where d_x and d_y bound the degrees of the input polynomials in x and y , respectively. We also prove that the total bitsize of the decomposition is in $\tilde{O}((d_x^2 d_y^3 + d_x d_y^4)\tau)$.

This work was done in collaboration with Fabrice Rouillier (Inria Rocquencourt). It is published in the research report [27] and submitted to a journal.

6.3.2. Numeric and Certified Isolation of the Singularities of the Projection of a Smooth Space Curve

Let a smooth real analytic curve embedded in \mathbb{R}^3 be defined as the solution of real analytic equations of the form $P(x, y, z) = Q(x, y, z) = 0$ or $P(x, y, z) = \frac{\partial P}{\partial z} = 0$. Our main objective is to describe its projection \mathcal{C} onto the (x, y) -plane. In general, the curve \mathcal{C} is not a regular submanifold of \mathbb{R}^2 and describing it requires to isolate the points of its singularity locus Σ . After describing the types of singularities that can arise under some assumptions on P and Q , we present a new method to isolate the points of Σ . We experimented our method on pairs of independent random polynomials (P, Q) and on pairs of random polynomials of the form $(P, \frac{\partial P}{\partial z})$ and got promising results [14].

On the same topic but with a different approach, we improved our research report [26] by including experimental data using SubdivisionSolver (see Section 5.2) and submitted this work to a journal.

6.3.3. Mechanical design of parallel robots

In collaboration with F. Rouillier, D. Chablat and our PhD student Ranjan Jha, we analyzed the singularities and the workspace of different families of robots.

The first result is a certified description of the workspace and the singularities of a Delta like family robot [16]. Workspace and joint space analysis are essential steps in describing the task and designing the control loop of the robot, respectively. This paper presents the descriptive analysis of a family of delta-like parallel robots by using algebraic tools to induce an estimation about the complexity in representing the singularities in the workspace and the joint space. A Gröbner based elimination is used to compute the singularities of the manipulator and a Cylindrical Algebraic Decomposition algorithm is used to study the workspace and the joint space. From these algebraic objects, we propose some certified three dimensional plotting describing the shape of workspace and of the joint space which will help the engineers or researchers to decide the most suited configuration of the manipulator they should use for a given task. Also, the different parameters associated with the complexity of the serial and parallel singularities are tabulated, which further enhance the selection of the different configurations of the manipulator by comparing the complexity of the singularity equations.

The second result is an algebraic method to check the singularity-free paths for parallel robots [15]. Trajectory planning is a critical step while programming the parallel manipulators in a robotic cell. The main problem arises when there exists a singular configuration between the two poses of the end-effectors while discretizing the path with a classical approach. This paper presents an algebraic method to check the feasibility of any given trajectories in the workspace. The solutions of the polynomial equations associated with the trajectories are projected in the joint space using Gröbner based elimination methods and the remaining equations are expressed in a parametric form where the articular variables are functions of time t unlike any numerical or discretization method. These formal computations allow to write the Jacobian of the manipulator as a function of time and to check if its determinant can vanish between two poses. Another benefit of this approach is to use a largest workspace with a more complex shape than a cube, cylinder or sphere. For the Orthoglide, a three degrees of freedom parallel robot, three different trajectories are used to illustrate this method.

6.3.4. Reflection through quadric mirror surfaces

We addressed the problem of finding the reflection point on quadric mirror surfaces, especially ellipsoid, paraboloid or hyperboloid of two sheets, of a light ray emanating from a 3D point source P_1 and going through another 3D point P_2 , the camera center of projection. We previously proposed a new algorithm for this problem, using a characterization of the reflection point as the tangential intersection point between the mirror and an ellipsoid with foci P_1 and P_2 . The computation of this tangential intersection point is based on our algorithm for the computation of the intersection of quadrics [5], [28]. Unfortunately, our new algorithm is not yet efficient in practice. This year, we made several improvements on this algorithm. First, we decreased from 11 to 4 the degree of a critical polynomial that we need to solve and whose solutions induce the coefficients in some other polynomials appearing later in the computations. Second, we implemented Descartes's algorithm for isolating the real roots of univariate polynomials in the case where the coefficients belong to extensions of \mathbb{Q} generated by at most two square roots. Furthermore, we are currently implementing the generalization of that algorithm when the coefficients belong to extensions of \mathbb{Q} generated by one root of an arbitrary polynomial. These undergoing improvements should allow us to compute more directly the wanted reflexion point, thus avoiding problematic approximations and making the overall algorithm tractable.

VERIDIS Project-Team

7. New Results

7.1. Automated and Interactive Theorem Proving

Participants: Gabor Alági, Haniel Barbosa, Jasmin Christian Blanchette, Martin Bromberger, Simon Cruanes, Pablo Dobal, Mathias Fleury, Pascal Fontaine, Maximilian Jaroschek, Marek Košta, Stephan Merz, Martin Riener, Thomas Sturm, Hernán Pablo Vanzetto, Uwe Waldmann, Daniel Wand, Christoph Weidenbach.

7.1.1. Combination of Satisfiability Procedures

Joint work with Christophe Ringeissen from the CASSIS project-team at Inria Nancy – Grand Est, and Paula Chocron, a student at the University of Buenos Aires.

A satisfiability problem is often expressed in a combination of theories, and a natural approach consists in solving the problem by combining the satisfiability procedures available for the component theories. This is the purpose of the combination method introduced by Nelson and Oppen. However, in its initial presentation, the Nelson-Oppen combination method requires the theories to be signature-disjoint and stably infinite (to ensure the existence of an infinite model). The design of a generic combination method for non-disjoint unions of theories is clearly a hard task, but it is worth exploring simple non-disjoint combinations that appear frequently in verification. An example is the case of shared sets, where sets are represented by unary predicates. Another example is the case of bridging functions between data structures and a target theory (e.g., a fragment of arithmetic).

We defined [24] a sound and complete combination procedure à la Nelson-Oppen for the theory of absolutely free data structures (including lists and trees) connected to another theory via bridging functions. This combination procedure has also been refined for standard interpretations. The resulting theory has a nice politeness property, enabling combinations with arbitrary decidable theories of elements. We also investigated [25] other theories amenable to similar combinations: this class includes the theory of equality, the theory of absolutely free data structures, and all the theories in between.

7.1.2. Adapting Real Quantifier Elimination Methods for Conflict Set Computation

The satisfiability problem in real closed fields is decidable. In the context of satisfiability modulo theories, the problem restricted to conjunctive sets of literals, that is, sets of polynomial constraints, is of particular importance. One of the central problems is the computation of good explanations of the unsatisfiability of such sets, i.e. obtaining a small subset of the input constraints whose conjunction is already unsatisfiable. We have adapted two commonly used real quantifier elimination methods, cylindrical algebraic decomposition and virtual substitution, to provide such conflict sets and demonstrate the performance of our method in practice [27].

7.1.3. Codatatypes and Corecursion

Joint work with Andrei Popescu and Dmitriy Traytel (Technische Universität München) and Andrew Reynolds (EPFL).

Datatypes and codatatypes are useful for specifying and reasoning about (possibly infinite) computational processes. The Isabelle/HOL proof assistant is being extended with flexible and convenient support for (co)datatypes and (co)recursive functions on them. We extended the emergent framework for (co)codatatypes with automatic generation of nonemptiness witnesses [22], nonemptiness being a proviso for introducing types in many logics, including Isabelle’s higher-order logic. As a theoretical step towards a definitional mechanism in Isabelle, we formalized a framework for defining corecursive functions safely, based on corecursion up-to and relational parametricity [21]. The end product is a general corecursor that allows corecursive (and even recursive) calls under “friendly” operations—an improvement over the inflexible syntactic criteria of systems such as Agda and Coq.

In a related line of work, we improved the automation of the SMT solver CVC4 by designing, implementing, and evaluating a combined decision procedure for datatypes and codatatypes [31]. The procedure decides universal problems and is composable via the Nelson–Oppen method, as implemented in SMT solvers. The decision procedure for (co)datatypes is useful both for proving and for model finding. We have commenced work on a higher-order model finder based on CVC4, called Nunchaku, that relies heavily on the decision procedure.

7.1.4. Analysis and Generation of Structured Proofs

Joint work with Sascha Böhme (QAware GmbH), Maximilian Haslbeck and Tobias Nipkow (Technische Universität München), Daniel Matichuk (NICTA), and Steffen J. Smolka (Cornell University).

Isabelle/HOL is probably the most widely used proof assistant besides Coq. The Archive of Formal Proofs is a vast collection of computer-checked proofs developed using Isabelle, containing nearly 65 000 lemmas. We performed an in-depth analysis of the archive, looking at various properties of the proof developments, including size, dependencies, and proof style [18]. This gives some insights into the nature of formal proofs.

In the context of the Sledgehammer bridge between automatic theorem provers and proof assistants, we designed a translation of machine-generated proofs into (semi-)intelligible Isabelle proofs that users can simply insert into their proof texts to discharge proof obligations [16]. While the output is designed for certifying the machine-generated proofs, it also has a pedagogical value: Unlike Isabelle’s automatic tactics, which are black boxes, the proofs delivered by Sledgehammer can be inspected and understood. The direct proofs also form a good basis for manual tuning.

7.1.5. Encoding Set-Theoretic Formulas in Many-Sorted First-Order Logic

TLA⁺ is a language for the formal specification of systems and algorithms whose first-order kernel is a variant of untyped Zermelo–Fraenkel set theory. Typical proof obligations that arise during the verification of TLA⁺ specifications mix reasoning about sets, functions, arithmetic, tuples, and records. Encoding such formulas in the input languages of standard first-order provers (SMT solvers or superposition-based provers for first-order logic) is paramount for obtaining satisfactory levels of automation. For set theory, the basic idea is to represent membership as an uninterpreted predicate for the backend provers, and to reduce set-theoretic expressions to this basic predicate. This is not straightforward for formulas involving set comprehension or for proofs that rely on extensionality for inferring equality of sets. Moreover, a full development of set-theoretic expressions may lead to large formulas that can overwhelm backend provers. We describe a technique that transforms set-theoretic formulas by successively applying rewriting and abstraction until a fixed point is reached. The technique is extended to handling functions, records, and tuples, and it is the kernel of the SMT backend of the TLA⁺ proof system (section 6.3). A paper describing our technique has been presented at the SETS workshop 2015 [46].

Although the approach was mainly intended to support proofs, we have also started work on adapting it for constructing models of formulas in set theory. Being able to construct (counter-)models can help users understand why proof attempts fail. During his internship, Glen Mével from ENS Rennes designed translation rules for a core fragment of TLA⁺ set theory. He validated them by using the finite model finding functionality of the SMT solver CVC4 for constructing models, with encouraging preliminary results.

7.1.6. Linear Constraints in Integer Arithmetic

We have investigated linear integer constraint solving. Many existing algorithms rely on solving the rational relaxation and transferring the results to an integer branch and bound approach. This algorithm eventually terminates due to the well-known a priori exponential bounds of an integer solution. De Moura and Jovanović proposed the first model-driven terminating algorithm where the termination relies on the structure of the problem itself but not on a priori bounds [62]. However, the algorithm contained some bugs, in particular it did not terminate. We fixed the bugs by introducing the notion of Weak Cooper elimination. Termination requires adding more rules to the algorithm and refining some existing ones [23].

7.1.7. Decidability of First-Order Clause Sets

Recursion is a necessary source for first-order undecidability of clause sets. If there are no cyclic, i.e., recursive definitions of predicates in such a clause set, (ordered) resolution terminates, showing decidability. In this work we present the first characterization of recursive clause sets enabling non-constant function symbols and depth increasing clauses but still preserving decidability. For this class called BDI (Bounded Depth Increase) we present a specialized superposition calculus. This work was published in the Journal of Logic and Computation [63]. Recursive clause sets also become decidable in the context of finite domain axioms. For this case we developed a new calculus that incorporates explicit partial model assumptions guiding the search [19].

7.1.8. Building Blocks for Automated Reasoning

There are automated reasoning building blocks shared between today's prime calculi for propositional logic (CDCL), propositional logic modulo theories (CDCL(T)), and first-order logic with equality (superposition). Underlying all calculi is a partial model assumption guiding inferences that are not redundant. Deciding the abstract redundancy notion is basically as difficult as the overall satisfiability problem for the respective logic, but for well-chosen partial model assumptions inferences can be guaranteed to be non-redundant at much lower cost. For example, for SAT it is possible to compute inferences in linear time [40] that are guaranteed to be non-redundant.

7.1.9. Beagle – A Hierarchic Superposition Prover

Joint work with Peter Baumgartner and Joshua Bax from NICTA, Canberra, Australia.

Hierarchic superposition is a calculus for automated reasoning in first-order logic extended by some background theory. In [20] we describe an implementation of hierarchic superposition within the Beagle theorem prover, and report on Beagle's performance on the TPTP problem library. Currently implemented background theories are linear integer and linear rational arithmetic. Beagle features new simplification rules for theory reasoning and implements calculus improvements like weak abstraction and determining (un)satisfiability w.r.t. quantification over finite integer domains.

7.1.10. Modal Tableau Systems with Blocking and Congruence Closure

Joint work with Renate A. Schmidt from the University of Manchester, UK.

For many common modal and description logics there are ways to avoid the explicit use of equality in a tableau calculus. For more expressive logics, e.g., with nominals as in hybrid modal logics and description logics, avoiding equality becomes harder, though, and for modal logics where the binary relations satisfy frame conditions expressible as first-order formulae with equality, explicit handling of equations is the easiest and sometimes the only known way to perform equality reasoning. In [32] we describe an approach for efficient handling of equality in tableau systems. We combine Smullyan-style tableaux with a congruence closure algorithm, and demonstrate that this method also permits the use of common blocking restrictions such as ancestor blocking.

7.1.11. Subtropical Real Root Finding

This research is motivated by a series of studies of Hopf bifurcations [60], [59] for reaction systems in chemistry and gene regulatory networks in systems biology. The relevant systems are originally given in terms of ordinary differential equations, for which Hopf bifurcations can be described algebraically [54], [74], [58], [57], typically resulting in one very large multivariate polynomial equation $f = 0$ subject to a few much simpler polynomial side conditions $g_1 > 0, \dots, g_n > 0$. For these algebraic systems one is interested in feasibility over the reals and, in the positive case, in at least one feasible point. It turns out that, generally, scientifically meaningful information can be obtained already by checking only the feasibility of $f = 0$, which is the focus of this project. For further details on the motivating problems, we refer to our earlier publications [72], [71], [56], [55].

With one of our models, viz. *Mitogen-activated protein kinase (MAPK)*, we obtain and solve polynomials of considerable size. Our currently largest instance `mapke5e6` contains 863,438 monomials in 10 variables. One of the variables occurs with degree 12, all other variables occur with degree 5. Such problem sizes are clearly beyond the scope of classical methods in symbolic computation. To give an impression, the size of an input file with `mapke5e6` in infix notation is 30 MB large. LaTeX-formatted printing of `mapke5e6` would fill more than 5000 pages in this report.

We have developed an incomplete but terminating algorithm for finding real roots of large multivariate polynomials [33]. The principal idea is to take an abstract view of the polynomial as the set of its exponent vectors supplemented with sign information on the corresponding coefficients. To that extent, our approach is quite similar to tropical algebraic geometry [73]. However, after our abstraction we do not consider tropical varieties but employ linear programming to determine certain suitable points in the Newton polytope, which somewhat resembles successful approaches to sum-of-square decompositions [67].

We have implemented our approach in Reduce [61] using direct function calls to the dynamic library of the LP solver Gurobi [48]. In practical computations on several hundred examples originating from the work within an interdisciplinary research group our method has failed due to its incompleteness in only 10 percent of the cases. The longest computation time observed was around 16 s for the above-mentioned `mapke5e6`. With a publication of our computational results in a physics journal, our research had considerable impact beyond computer science [17].

7.1.12. Standard Answers for Virtual Substitution

Joint work with A. Dolzmann from Leibniz-Zentrum für Informatik in Saarbrücken, Germany.

We consider existential problems over the reals. Extended quantifier elimination generalizes the concept of regular quantifier elimination by additionally providing answers which are descriptions of possible assignments for the quantified variables. Implementations of extended quantifier elimination via virtual substitution have been successfully applied to various problems in science and engineering.

So far, the answers produced by these implementations included infinitesimal and infinite numbers, which are hard to interpret in practice. This has been explicitly criticized in the scientific literature. In our article [44], we introduce a complete post-processing procedure to convert, for fixed values of parameters, all answers into standard real numbers. We furthermore demonstrate the successful application of an implementation of our method within Redlog to a number of extended quantifier elimination problems from the scientific literature including computational geometry, motion planning, bifurcation analysis for models of genetic circuits and for mass action, and sizing of electrical networks.

7.1.13. A Generalized Framework for Virtual Substitution

We generalize the framework of virtual substitution for real quantifier elimination to arbitrary but bounded degrees [45]. We make explicit the representation of test points in elimination sets using roots of parametric univariate polynomials described by Thom codes. Our approach follows an early suggestion by Weispfenning, which has never been carried out explicitly.

We give necessary and sufficient conditions for the existence of a root with a given test point representation. These conditions are used to rule out redundant test points. Our encoding allows us to distinguish between test points that represent lower bounds and test points representing upper bounds of a satisfying interval for a given input formula. Furthermore, we show how to reduce the size of elimination sets by generalizing a well-known idea from linear virtual substitution, namely to consider only test points representing lower bounds of a satisfying interval.

Our framework relies on some external algorithm \mathcal{A} , which is used to eliminate a single existential quantifier from a finite set of generic formulas. The existence of \mathcal{A} is guaranteed by the fact that \mathbb{R} admits quantifier elimination. We briefly refer to experiments which compared the performance of our framework—when Cylindrical Algebraic Decomposition is used as the external algorithm—to other quantifier elimination algorithms. Unfortunately, our approach is not yet able to compete with other state-of-the-art quantifier

elimination algorithms. However, currently ongoing research suggests the possibility for drastic improvements in practice. Investigating this is left for future work.

7.2. Formal Methods for Developing Algorithms and Systems

Participants: Manamiary Andriamiarina, Noran Azmy, Gabriel Corona, Marie Duflot-Kremer, Marion Guthmuller, Souad Kherroubi, Dominique Méry, Stephan Merz, Martin Quinson, Christoph Weidenbach.

7.2.1. Incremental Development of Distributed Algorithms

Joint work with Mike Poppleton, University of Southampton, UK, and with Neeraj Kumar Singh from the Department of Computing and Software, McMaster University, Hamilton, Canada.

The development of distributed algorithms and, more generally, of distributed systems, is a complex, delicate, and challenging process. The approach based on refinement helps to gain formality by using a proof assistant, and proposes to apply a design methodology that starts from the most abstract model and leads, in an incremental way, to the most concrete model, for producing a distributed solution. Our work helps formalizing pre-existing algorithms, developing new algorithms, as well as developing models for distributed systems.

More concretely, we aim at an integration of the correct-by-construction refinement-based approach for distributed algorithms. Our main results during 2015 are:

- An integrated formal method for verification of liveness properties in distributed systems is introduced [43], and the verification of a self-stabilizing leader election protocol for population protocols illustrates the proposed methodology.
- Manamiary Andriamiarina completed his PhD, illustrating a method for developing distributed algorithms based on a combination of Event-B and fragment of temporal logic TLA.
- The methodology has been applied to take into account resilience in distributed systems. We describe a fully mechanized proof of correctness of self- \star systems [42] along with an interesting case study related to P2P-based self-healing protocols.

7.2.2. Modeling Medical Devices

Joint work with Neeraj Kumar Singh from the Department of Computing and Software, McMaster University, Hamilton, Canada.

Formal modeling techniques and tools have attained sufficient maturity for formalizing highly critical systems in view of improving their quality and reliability, and the development of such methods has attracted the interest of industrial partners and academic research institutions. Building high quality and zero-defect medical software-based devices is a particular domain where formal modelling techniques can be applied effectively. Medical devices are very prone to showing unexpected system behaviour in operation when traditional methods are used for system testing. Device-related problems have been responsible for a large number of serious injuries. Officials of the US Food and Drug Administration (FDA) found that many deaths and injuries related to these devices are caused by flaws in product design and engineering. Cardiac pacemakers and implantable cardioverter-defibrillators (ICDs) are among the most critical medical devices and require closed-loop modelling (integrated system and environment modelling) for verification purposes before obtaining a certificate from the certification bodies.

Clinical guidelines systematically assist practitioners in providing appropriate health care in specific clinical circumstances. Today, a significant number of guidelines and protocols are lacking in quality. Indeed, ambiguity and incompleteness are likely anomalies in medical practice. The analysis of guidelines using formal methods is a promising approach for improving them.

Analyzing requirements is a major challenge in the area of safety-critical software, where the quality of requirements is an important issue for building a dependable critical system. Many projects fail due to lack of understanding of user needs, missing functional and non-functional system requirements, inadequate methods and tools, and inconsistent system specifications. This often results from the poor quality of system requirements. Based on our experience and knowledge, an environment model has been recognized to be a promising approach to support the requirements engineering to validate a system specification. It is crucial to get an approval and feedback at an early stage of the system development to guarantee the completeness and correctness of the requirements. In [29], we propose a method for analyzing the system requirements using closed-loop modelling technique. The closed-loop model is an integration of system model and environment model, where both the system and environment models are formalized using formal techniques. Formal verification of this closed-loop model helps to identify hidden or missing system requirements and peculiar behaviours, which are not covered earlier during requirements elicitation process. Moreover, the environment model assists in the construction, clarification, and validation of a given system requirements.

7.2.3. Verification of the Pastry routing protocol

In his PhD thesis at Saarbrücken University in 2013, Tianxiang Lu had studied the routing protocol of the Pastry algorithm [69] for maintaining a distributed hash table in a peer-to-peer network. He had discovered several problems in the published algorithm and proposed a modification of the protocol, together with a correctness proof under the hypothesis that no node ever disconnects. The proof had been checked using TLAPS, but it made many assumptions on the underlying data structures that were left unchecked. In particular, support for (modulus) arithmetic in TLAPS was too weak at the time when the proof was written.

As part of her PhD thesis, Noran Azmy studied the assumptions that had been left unproved, and found that several of them were not valid. As a consequence, she was able to find a counter-example to one of the invariants underlying the correctness proof. She corrected the assumptions, proved all of the ones that were needed for the proof using the current version of TLAPS, and also introduced higher-level abstractions that allowed her to rewrite the specification and the correctness proof of the routing protocol in a way that avoids low-level arithmetic reasoning throughout the proof. As a result, she obtained a complete machine-checked proof of Lu's variant of Pastry, still under the assumption that no node leaves the network. A paper describing the result is being submitted.

7.2.4. Proof of Determinacy of PharOS

Joint work with Selma Azaiez and Matthieu Lemerre (CEA Saclay), and Damien Doligez (Inria Paris).

The main contribution of our team to the ADN4SE project (section 8.1), in cooperation with colleagues from CEA, was to write a high-level specification of the real-time operating system PharOS in the TLA⁺ language, and to prove a determinacy property of the model using TLAPS. Roughly speaking, determinacy means that the sequence of local states of each process during a computation does not depend on the order in which processes are scheduled, as long as there are no missed deadlines. This property simplifies the analysis and verification of programs that run on PharOS. It relies on the fact that every instruction is associated with a time window of execution, and a message can only be received by an instruction if the earliest possible execution time of that instruction is later than the latest possible execution time of the instruction sending the message. The model and proof are based on Lemerre et al. [65]. However, the underlying assumptions are made fully explicit in the formal model, and the proof is carried out in assertional rather than behavioral style. The proof was completed in 2015, and a paper describing the result is being submitted.

7.2.5. Formal Development of Component Semantics in B

Joint work with David Déharbe of Universidade Federal do Rio Grande de Norte (UFRN), Brazil.

We develop a formal model in Isabelle/HOL of the behavioral semantics of software components designed with the B method. We formalize semantic objects, based on labeled transition systems, notions of internal and externally visible behavior, and simulation. In particular, we study a variant of simulation that corresponds to refinement in the B method. We also formally represent the composition of components in the B method.

This work was presented at an invited talk at FACS 2015 in Rio de Janeiro, and an article will be published in LNCS.

7.2.6. Analysis of Distributed Legacy Applications

SimGrid is a toolkit for the study of Large-Scale Distributed Systems. It contains both a simulator with sound and validated performance models for the network, CPUs, and disks, but also an explicit model checker exploring all possible message interleavings in the application, and searching for states violating some properties specified by the user.

We recently added the ability to assess liveness properties over arbitrary and legacy codes, thanks to a system-level introspection tool that provides a detailed view of the running application to the model checker. This can for example be leveraged to verify both safety and liveness properties, on arbitrary MPI code written in C, C++ or Fortran. This work has been published in the Workshop on Formal Approaches to Parallel and Distributed Systems (4PAD) [26], while the full details appear in Guthmuller's PhD thesis [12].

In his master project, Gabriel Rodrigues Santos investigated the feasibility of implementing algorithms for statistical model checking within SimGrid. The basic idea is to sample sufficiently many executions of a program, based on probabilistic parameters associated with the execution platform, for quantifying correctness and reliability properties. By construction, the answers obtained in this way are not exact, but their imprecision can be bounded by an interval of confidence. The results are very encouraging, and we intend to pursue this approach in further work.

7.2.7. Evaluating and Verifying Probabilistic Systems

Joint work with colleagues at ENS Cachan, University Paris Est Créteil, and Ecole Centrale Paris.

Since its introduction in the 1980s, model checking has become a prominent technique for the verification of complex systems. The aim was to decide whether or not a system fulfills its specification. With the rise of probabilistic systems, new techniques have been designed to verify this new type of systems, and appropriate logics have been proposed to describe more subtle properties to be verified. However, some characteristics of such systems fall outside the scope of model checking. In particular, it is often of interest not to decide whether a property is satisfied but how well the system performs with respect to a certain measure. We have designed a statistical tool for tackling both performance and verification issues. Following several conference talks, two journal papers have been published. The first one [14] presents the approach in details together with illustrative applications to flexible manufacturing systems, and to the study of a biological mechanism known as circadian clock. The second one [15] focuses on biological applications, and more precisely the use of statistical model checking to detect and measure several indicators of oscillating biological systems.

ACUMES Team

6. New Results

6.1. Mathematical analysis and control of macroscopic traffic flow models

6.1.1. Vehicular traffic

Participants: Guillaume Costeseque, Paola Goatin, Christophe Chalons [UVST], Simone Göttlich [U Mannheim, Germany], Jérôme Härrı [EURECOM], Oliver Kolb [U Mannheim, Germany], Sosina Mengistu-Gashaw [EURECOM], Francesco Rossi [U Aix-Marseille], Stefano Villa [U Milano-Bicocca].

In collaboration with the University of Mannheim and in the framework of the PHC Procope project "Transport Networks Modeling and Analysis", we studied how to manage variable speed limits combined with coordinated ramp metering within the framework of the LWR network model. Following a "first discretize then optimize" approach, we derived the first order optimality system and explained the switch of speeds at certain fixed points in time and the boundary control for the ramp metering as well. Sequential quadratic programming methods are used to solve the control problem numerically. For application purposes, we present experimental setups where variable speed limits are used as a traffic guidance system to avoid traffic jams on highway interchanges and on-ramps, see [35].

The thesis of S. Mengistu-Gashaw, funded by the Labex UCN@Sophia (<http://ucnlab.eu/>) and co-supervised by P. Goatin and J. Härrı, is devoted to understanding and modeling mobility characteristics of scooters and motorcycles for user-centric ITS application. We are currently developing a macroscopic model for heterogeneous traffic including car and motorcycles.

A new traffic flow model has been designed in [44] for taking into account the multiclass and multilane features of real traffic. This model is based on a system of coupled Hamilton-Jacobi PDEs for an appropriate choice of framework that mixes spatial and Lagrangian coordinates. The coupling conditions emerge from the moving bottleneck theory that has been developed in the traffic flow literature several years ago but for which a real mathematical sound basis lacked. Very recently, there were some new results dealing with the existence of a solution under suitable assumptions. However, these results were set for the hyperbolic conservation law in Eulerian coordinates and they are not straightforward to extend to Hamilton-Jacobi equations in different coordinates. Despite that the well-posedness of the problem is still an open problem, a numerical method is developed and it takes advantage of the classical representation formula available for HJ PDEs. This numerical scheme has been proved to provide good qualitative results.

In collaboration with F. Rossi, we proved existence and uniqueness of solutions to a transport equation modelling vehicular traffic in which the velocity field depends non-locally on the downstream traffic density via a discontinuous anisotropic kernel. The result is obtained recasting the problem in the space of probability measures equipped with the ∞ -Wasserstein distance. We also show convergence of solutions of a finite dimensional system, which provide a particle method to approximate the solutions to the original problem. See, [45].

Finally, the internship of S. Villa, co-supervised by M. Garavello (U Milano-Bicocca), was devoted to the analytical and numerical study of the Aw-Rascle-Zhang model with moving bottleneck. Two Riemann Solver have been proposed, and two numerical strategies have been developed. A journal article is in preparation in collaboration with C. Chalons.

6.1.2. Crowd motion

Participants: Paola Goatin, Matthias Mimault.

M. Mimault defended his PhD on December 14th, 2015. The last part of his thesis was devoted to the numerical study of scalar conservation laws with non-local flow in two space dimensions. These equations are meant to model crowd motion, where the movement direction of each pedestrian depends on a weighted mean of the crowd density around him. In particular, he implemented a finite volume numerical scheme which has been used for flow optimization purposes: he applied the adjoint method to compute the gradient for the evacuation time minimization depending on the initial crowd distribution.

6.2. Characterization of model uncertainty for turbulent flows

Participants: Régis Duvigneau, Jérémie Labroquère, Emmanuel Guilmineau [CNRS ECN, Nantes], Marianna Brazza [CNRS IMFT, Toulouse], Mathieu Szubert [CNRS IMFT, Toulouse].

The uncertainty related to turbulence modeling is still a bottleneck in realistic flows simulation. Therefore, some studies have been conducted to quantify this uncertainty for two problems in which turbulence plays a critical role. Firstly, the impact of the model choice has been estimated in the case of a massively detached flow over a 2D backward facing step including an oscillatory active control device, whose parameters are optimized [41], [34]. Secondly, the influence of the transition point location has been investigated, in the case of the 3D flow around a bluff-body, using models ranging from RANS to DES models [40].

6.3. Sensitivity analysis for unsteady flows

Participants: Régis Duvigneau, Dominique Pelletier [Ecole Polytechnique Montreal], Alexander Hay [Ecole Polytechnique Montreal].

Although sensitivity analysis is now commonly used for steady systems, usually on the basis of the adjoint equation method, the application to unsteady problems is still tedious, due to the backward time integration required. Therefore, an alternative approach, namely the sensitivity equation method, has been studied in the framework of the compressible Navier-Stokes equations. A continuous version has preferred to the discrete one for its flexibility and easier implementation. The proposed approach has been verified on several problems of increasing difficulty and the computational efficiency quantified [42].

6.4. Optimization accounting for experimental and numerical uncertainties

Participants: Régis Duvigneau, Olivier Le Maitre [CNRS LIMSI, Orsay], Matthieu Sacher [Ecole Navale, Brest], Alban Leroyer [ECN, Nantes], Patrick Queutey [CNRS ECN, Nantes].

Optimization of real-life applications requires to account for the uncertainties arising during the performance evaluation procedure, that could be either experimental or numerical. A Gaussian-Process based optimization algorithm has been proposed to efficiently determine the global optimum in presence of noise, whose amplitude can be user-defined or inferred from observations. The method has been applied to two very different problems related to performance optimization in sport.

The first case corresponds to the optimization of the shape of a racing kayak. The performance is estimated by coupling Newton's law with Navier-Stokes equations to compute the kayak velocity from the effort of the athlete, considered as input. The proposed method has been used here to filter the noise arising from the numerical simulation.

The second case corresponds to the optimization of a sail trimming, whose performance can be estimated either experimentally in a wind tunnel, or numerically by solving a fluid-structure interaction problem. In the former case, uncertainty has been estimated according to measurements accuracy, while in the latter case the numerical noise has been inferred from a set of observations collected during the optimization.

6.5. High-order numerical schemes for convection-dominated problems

Participants: Régis Duvigneau, Asma Gdhami [ENIT, Tunisia].

The use of high-order numerical schemes is necessary to reduce numerical diffusion in simulations, maintain a reasonable computational time for 3D problems, estimate accurately uncertainties or sensitivities, etc. Consequently, we work to develop high-order numerical schemes for the applications targeted by the team, in particular for convection-dominated problems. More precisely, we intend to include in a unified framework, based on Discontinuous Galerkin approximations, numerical methods accounting for complex geometries (isogeometric methods), uncertainty propagation (high-dimensional cubature) and sensitivity analysis.

6.6. Validation of time dependent diffusion approaches for activated and inhibited cell sheet closure

Participants: Abderrahmane Habbal, Hélène Barelli [Univ. Nice Sophia Antipolis, CNRS, IPMC], Grégoire Malandain [Inria, EPI Morpheme], Boutheina Yahyaoui [PhD, LAMSIN, Univ. Tunis Al Manar], Mekki Ayadi [LAMSIN, Univ. Tunis Al Manar].

We have studied in [21] five MDCK cell monolayer assays in a reference, activated and inhibited migration conditions. Modulo the inherent variability of biological assays, we have shown that in the assay where migration was not exogeneously activated or inhibited, the wound velocity was constant and the Fisher-KPP equation was able to accurately predict, until the final closure of the wound, the evolution of the wound area, the mean velocity of the cell front, and the time at which the closure occurred. When activated or inhibited, the F-KPP equation with constant parameters was unable to reproduce the observed biological cell sheet behavior. We modify the original equation, making the diffusion and proliferation parameters time dependent, following a sigmoid profile. We then set up an optimization loop to identify the sigmoids parameters, by computing a classical error indicator (difference between a computed density and the observed one, obtained through image processing) as done in the cited reference. We then obtain results which convincingly show that our approach is efficient : in both cases, inhibited and activated, the time varying identified parameters allow us to accurately predict until the final closure the evolution of the wound area.

6.7. Game strategies for joint data completion and parameter identification

Participants: Abderrahmane Habbal, Rabeb Chamekh [PhD, LAMSIN, Univ. Tunis Al Manar], Moez Kallel [LAMSIN, Univ. Tunis Al Manar], Nejib Zemzemi [Inria Bordeaux, EPI CARMEN].

We have demonstrated in previous works [22], [23] that Nash game approaches are efficient to tackle ill-posedness for linear second order elliptic Cauchy problems. We next developed a mathematical formulation for the linear elasticity model. The reconstruction is based on data completion and material identification, making it a harsh ill posed inverse problem. Up to now, we have obtained successful results for the Lamé parameter recovery in linear elasticity, using the so-called Kohn and Vogelius functional. Simultaneous data completion and parameter identification is under investigation.

6.8. Revised definition of the Multiple Gradient Descent Algorithm (MGDA)

Participant: Jean-Antoine Désidéri.

The Multiple Gradient Descent Algorithm (MGDA) had been defined originally to identify a descent direction common to a set of gradient vectors. According to a completely general principle, the direction is opposite to the vector of minimum Euclidean norm in the convex hull of the gradients. The Euclidean norm is defined via a general scalar product in \mathbf{R}^n . From a theoretical viewpoint, the notion of Pareto-stationarity had been introduced and it was established that if a point is Pareto-optimal and if the objective functions are locally differentiable and convex, then the point is Pareto-stationary. From a computational viewpoint, the descent direction can be determined as the solution of a Quadratic-Programming (QP) formulation. However, when the gradients are linearly independent a direct construction via a Gram-Schmidt orthogonalization process was preferred. We have now generalized the orthogonalization process by the introduction of a hierarchical strategy in the ordering of the subfamily of gradients utilized to construct the orthogonal basis. This strategy aims at making the (multi-dimensional) cone associated with the convex hull of the subfamily as large as possible. As

a result, in the case of linearly-dependent gradients, the orthogonalization process not only provides a basis of the spanned subspace, but the subfamily is selected such that its convex hull is also very representative of a large cone, encompassing in the most favorable cases all the given gradients. By this change in the definition of the algorithm, we were able to reformulate the QP formulation, now stated in a suitable basis, in a way that is well-suited for the treatment of cases where the number of gradients exceeds, possibly vastly, the dimension of the vector space. This revision makes the algorithm much more general and robust [43].

6.9. Multi-point optimization of a time-periodic system of pulsating jets

Participants: Jean-Antoine Désidéri, Régis Duvigneau.

A multi-point optimization exercise governed by the time-dependent compressible Navier-Stokes equations has been solved based on the sensitivity analysis (see above). A system was considered consisting of three pulsating jets acting on a flat-plate boundary layer. As it is well-known, the flow mixing by the jets has the effect of reducing the drag, as this was confirmed by the simulation of the flow in the somewhat arbitrary initial setting of the jets. Then, positions and pulsation frequencies of the jets have been maintained fixed, while their amplitudes and phases, six parameters in total, have been optimized to minimize the drag force. The finite-volume simulation of the time-periodic flow provides the drag force as a function of time over a large number of timesteps (800 for an accurate description of a period). The sensitivity analysis simultaneously provides the derivatives of drag with respect to the six design parameters. These derivatives were averaged over 20 distinct time-intervals, thus yielding 20 averaged gradient vectors of dimension 6. The MGDA was then used to define a descent direction common to the 20 vectors, a descent step was applied to the design parameters, and the process was continued iteratively.

The experiment confirmed the possibility to reduce the drag force at all times of the period, and not only in the average. In contrast, using the average gradient to define the direction of search resulted in a more important reduction of the average drag but at the cost of an increase of drag in a critical portion of the time period. Hence our optimization algorithm is more versatile and powerful than one aiming at minimizing purely statistical functions obtained by time averaging. We also demonstrated the possibility to optimize over a subinterval of the time interval.

6.10. Quasi-Riemannian approach to constrained optimization

Participants: Didier Bailly [Research Engineer, ONERA Department of Applied Aerodynamics, Meudon], Gérald Carrier [Research Engineer, ONERA Department of Applied Aerodynamics, Meudon], Jean-Antoine Désidéri.

In differentiable optimization, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is one of the most efficient methods for unconstrained problems. Besides function values, it only requires the specification of the gradient. An approximate Hessian is calculated by successive approximations as part of the iteration, using rank-1 correction matrices. As a result, the iteration has superlinear convergence : when minimizing a quadratic function in n variables, if the one-dimensional minimizations in the calculated directions of search are done exactly, the Hessian matrix approximation is exact after n iterations, and from this, the iteration identifies to Newton's iteration, and produces the exact local optimum in only one additional iteration ($n + 1$ in total).

However the BFGS method does not extend to constrained problems very simply. Following Gabay [95] and other authors, Chunhong Qi *et al* [128] have proposed a "Riemannian" variant, RBFGS that indeed incorporates equality constraints in the formulation and actually demonstrates superior convergence rates for problems with a large number of variables. However these Riemannian formulations are non trivial to implement since they require procedures implementing non-trivial differential-geometry operators ('retraction' and 'metric transport') to be developed. In their paper, they assume a formal expression of the constraint to be known. But, in PDE-constrained optimization, many constraints are functional, and it is not clear how can the metric transport operator in particular can be defined.

We are investigating how can a quasi-Riemannian method can be defined based on the sole definition of evaluation procedures for the gradients. By condensing all the equality constraints in one, a purely-explicit approximate retraction operator has been defined that yields a point whose distance to the constraint surface is fourth-order at least. The associated transport operator is currently being examined formally. These techniques will be experimented in the context of constrained optimum-shape design in aerodynamics.

6.11. Unstructured mesh adaptation using an adjoint-based sensor

Participants: Sébastien Bourasseau [Doctoral student, ONERA Department of Computational Fluid Dynamics, Châtillon], Jacques Peter [Research Engineer, ONERA Department of Computational Fluid Dynamics, Châtillon], Jean-Antoine Désidéri.

Mesh adaptation is a powerful tool to obtain accurate aerodynamic simulations at limited cost. When the simulation is aimed at the accurate calculation of aerodynamic outputs (forces, moments) goal-oriented methods based on the adjoint vector of the output of interest are often advocated. The calculation of the total derivative dJ/dX of the aerodynamic function of interest, J , with respect to the volume mesh coordinates, X , has been extended to the case of an unstructured grid. The software developments have been validated for inviscid and laminar viscous flows, and implemented in the ONERA code (elsA). Then a local sensor θ based on dJ/dX was devised to identify areas where the location of the volume mesh nodes has a strong impact on the evaluation of the output J . The sensor has been shown to be adequate in different flow regimes (subsonic, transonic, supersonic), for internal (blade and nozzle) and external (airfoils, wings) aerodynamic configurations. The proposed method has been compared to a well-known goal-oriented method (Darmofal and Venditti, 2001) and to a feature-based method; it yields comparable results at lower cost in simple configurations. A publication is currently subject to minor revisions.

6.12. Multi-fidelity surrogate modeling with application to the optimization of nanophotonic devices

Participants: Cédric Durantin [Doctoral student, CEA LETI Grenoble], Alain Glière [Research Engineer, CEA LETI Grenoble], Jean-Antoine Désidéri.

Multiple models of a physical phenomenon are sometimes available with different levels of approximation, the high fidelity model being more demanding in terms of computational time than the coarse approximation. In this context, including information from the lower fidelity model to build a surrogate model is desirable. A new multi-fidelity metamodeling method, based on Radial Basis Function, the co-RBF, is proposed. The new method is compared with the classical co-kriging on two analytical benchmarks and a realistic validation test, namely the design of a miniaturized photoacoustic gas sensor. The co-RBF method brings better results on high dimensional problem and could be considered as an alternative to co-kriging for multi-fidelity metamodeling.

APICS Project-Team

6. New Results

6.1. Inverse problems for Poisson-Laplace equations

Participants: Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Konstantinos Mavreas, Christos Papageorgakis, Dmitry Ponomarev.

This section is concerned with inverse problems for 3-D Poisson-Laplace equations, among which source recovery issues. Though the geometrical settings differ in Sections 6.1.1 and 6.1.2, the characterization of silent sources (those giving rise to a vanishing field) is one common problem to both which has been resolved in the magnetization setup [33].

6.1.1. Inverse problems in medical imaging

This work is conducted in collaboration with Jean-Paul Marmorat and Nicolas Schnitzler, together with Maureen Clerc and Théo Papadopoulo from the Athena EPI.

In 3-D, functional or clinical active regions in the cortex are often modeled by pointwise sources that have to be localized from measurements taken by electrodes on the scalp of an electrical potential satisfying a Laplace equation (EEG, electroencephalography). In the works [38][5] on the behavior of poles in best rational approximants of fixed degree to functions with branch points, it was shown how to proceed via best rational approximation on a sequence of 2-D disks cut along the inner sphere, for the case where there are finitely many sources (see Section 4.2).

In this connection, a dedicated software FindSources3D (see Section 3.4.2) is being developed, in collaboration with the team Athena and the CMA. We continued this year algorithmic developments, prompted by a fruitful collaboration with the firm BESA, namely automatic detection of the number of sources (which was left to the user until recently). It appears that, in the rational approximation step, *multiple* poles possess a nice behavior with respect to branched singularities. This is due to the very physical assumptions on the model (for EEG data, one should consider *triple* poles). Though numerically observed in [7], there is no mathematical justification so far why multiple poles generate such strong accumulation of the poles of the approximants. This intriguing property, however, is definitely helping source recovery. It is used in order to automatically estimate the “most plausible” number of sources (numerically: up to 3, at the moment). Further, a modular and ergonomic platform version of the software is under development.

In connection with these and other brain exploration modalities like electrical impedance tomography (EIT), we are now studying conductivity estimation problems. This is the topic of the PhD research work of C. Papageorgakis (co-advised with the Athena project-team and BESA GmbH). In layered models, it concerns the estimation of the conductivity of the skull (intermediate layer). Indeed, the skull was assumed until now to have a given isotropic constant conductivity, whose value can differ from one individual to another. A preliminary issue in this direction is: can we uniquely recover and estimate a single-valued skull conductivity from one EEG recording? This has been established in the spherical setting when the sources are known, see [17]. Situations where sources are only partially known and the geometry is more realistic than a sphere are currently under study. When the sources are unknown, we should look for more data (additional clinical and/or functional EEG, EIT, ...) that could be incorporated in order to recover both the sources locations and the skull conductivity. Furthermore, while the skull essentially consists of hard bone part that may be assumed to have constant electrical conductivity, it also contains spongy bone compartments. These two distinct parts of the skull possess quite different conductivities. The influence of that second value on the overall model is now being studied [19].

6.1.2. Inverse magnetization issues in the thin-plate framework

This work is carried out in the framework of the “équipe associée Inria” IMPINGE, comprising Eduardo Andrade Lima and Benjamin Weiss from the Earth Sciences department at MIT (Boston, USA) and Douglas Hardin, Michael Northington and Edward Saff from the Mathematics department at Vanderbilt University (Nashville, USA).

Localizing magnetic sources from measurements of the magnetic field away from the support of the magnetization has been the fundamental issue under investigation by IMPINGE. The goal was to determine magnetic properties of rock samples (*e.g.* meteorites or stalactites) from fine field measurements close to the sample that can nowadays be obtained using SQUIDs (superconducting quantum interference devices). Currently, rock samples are cut into thin slabs and the magnetization distribution is considered to lie in a plane, which makes for a somewhat less indeterminate framework than EEG as regards inverse problems because “less” magnetizations can produce the same field (for the slab has no inner volume). Note however that EEG data consist of values of the normal current and of the associated potential, while in the present setting only values of the normal magnetic field are measured.

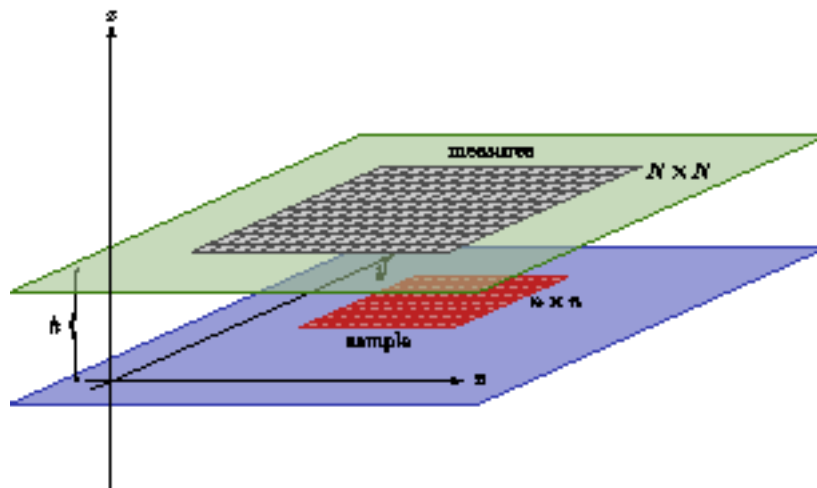


Figure 3. Schematic view of the experimental setup

Figure 3 presents a schematic view of the experimental setup: the sample lie on a horizontal plane at height 0 and its support is included in a rectangle. The vertical component B_3 of the field produced by the sample is measured on points of a horizontal $N \times N$ rectangular grid at height h .

Over the previous years, we mainly focused on developing techniques to recover magnetizations with rather sparse support. To this end, we set up a heuristic procedure to recover sparse magnetizations, based on iterative truncation of the support of the recovered magnetization. In this heuristics, magnetizations were represented by dipoles placed at the points of a regular rectangular $n \times n$, which seemed general enough a model class to correctly approximate the magnetizations commonly encountered in samples.

The procedure turned out to be poor when trying to recover the magnetization itself, due to the severe ill-posedness of the problem and the unexpected existence of magnetizations that produce almost no field at the height where measurements are performed, although the corresponding magnetic distributions strongly differ from truly silent distributions. Nevertheless, whenever the support could be significantly shrunk while keeping the error small (*i.e.*, explaining the data satisfactorily), estimates of the net moment so far, based on the dipolar model obtained by inversion, have been good.

This suggests that recovering the net moment and recovering the magnetization are rather different problems, the first one being less ill-posed than the second. Although the information provided by the net moment of the sample seems to be much weaker than knowing the full magnetic distribution, its importance has been emphasized by the geophysicists at MIT for at least three reasons:

- It yields important geological information on the sample in particular to estimate the magnitude of the ambient magnetic field at the time the rock was formed.
- It can be estimated independently to some extent, using a magnetometer, thereby allowing one to cross-validate the approach.
- From a computation point of view, knowledge of the net moment should lead to numerically stable reconstruction of an equivalent unidirectional magnetization. The support of the latter would provide us with valuable information to test for unidirectionality of the true magnetization, which is an important question to physicists in connection with rocks history and formation.

This year, we addressed the problem of directly recovering the net moment, without recourse to full inversion. Indeed, the latter is rather inefficient as it requires using a cluster and even then, for some samples, days of evaluation in order to obtain only a coarse estimate of the net moment. This research effort led us to investigate three different and complementary approaches.

First, we improved over Fourier based techniques previously designed by reformulating the problem with the help of the Kelvin transform. This gave us an asymptotic expansion of the net moment involving, at the first order, the integrals $\iint B_3(x, y, h) dx dy$, $\iint x B_3(x, y, h) dx dy$ and $\iint y B_3(x, y, h) dx dy$, computed on a disc with large radius. Although the method is promising, the computations are quite involved and we did not manage yet to obtain higher-order terms. This is a part of D. Ponomarev PhD work.

In parallel, and based on the results obtained with Fourier transform, we investigated a second approach, consisting in directly computing asymptotic expansions of the above integrals, on several domains (namely, the 2-D balls of radius R for the 1, 2 and ∞ norm, that are squares, disks, diamonds). In all cases, we get

$$\iint x B_3(x, y, h) dx dy = \alpha \langle m_1 \rangle + \beta (\langle t_1 m_3 \rangle - h \langle m_1 \rangle) / R + \mathcal{O}(1/R^3),$$

where $\langle m_1 \rangle$ is the moment of the first component m_1 of the magnetization and $\langle t_1 m_3 \rangle$ is the first moment of m_3 with respect to the first variable. The constants α and β depend on the domain where the integral is computed. Therefore, an appropriate linear combination of the integrals computed on the different domains allows us to compute $\langle m_1 \rangle$ with an accuracy of $\mathcal{O}(1/R^3)$. Similar results are obtained for $\langle m_2 \rangle$ and $\langle m_3 \rangle$ with the other integrals. Preliminary numerical experiments confirm the practical usability of these formulas in order to recover the moment of magnetizations. A research report is currently being written to sum up these results.

Finally, a third more ambitious approach has been investigated. As an attempt to generalize the previous expansions, our initial question was: given measurement of B_3 , say on a square, find a function $\phi(x, y)$ such that $\iint \phi(x, y) B_3(x, y) dx dy$ is the best possible estimate of the net moment components $\langle m_i \rangle$ ($i = 1, 2, 3$). This problem does not admit a solution because, for any $\epsilon > 0$, there exists a function ϕ_ϵ allowing to estimate the moment with an error bounded by ϵ . However, when ϵ tends to zero, the function ϕ_ϵ is expected to have strong oscillations, which hinders an accurate computation of $\iint \phi(x, y) B_3(x, y) dx dy$ since B_3 is only known on a discrete grid of points. We therefore expressed the problem as a bounded extremal problem (see Section 3.3.1): to find the best ϕ_ϵ (with the smallest possible error value ϵ) under the constraint that $\|\nabla \phi_\epsilon\|_2 \leq M$. Here, M is a user-defined parameter. We proved theoretical results regarding this bounded extremal problem (existence and uniqueness of a solution, characterization of its solution as a solution of integro-differential equation) and we are currently designing a numerical procedure to compute it. An article on this topic is in preparation.

Still in the course of D. Ponomarev's PhD research, the study of a 2D spectral problem for the truncated Poisson operator in planar geometry has been pursued. It is a simplified formulation of the relation between the magnetization and the magnetic potential (of which the magnetic field is the gradient) and is expected to produce an efficient representation basis (the eigenfunctions of the magnetization-to-field operator). This is a long-standing problem. Noteworthy properties of solutions have been obtained through connections with other spectral problems and asymptotic reductions for large and small values of the main parameters (distance h from the measurement plane to the sample support and sample support size), yielding approximate solutions by means simpler integral equations and ODEs.

The year 2015 was the last of our "équipe associée" IMPINGE with the MIT and Vanderbilt University. The final report is available on the web page of the associate team ⁰. This collaboration is currently supported in part by a MIT-France seed funding from the US side, and we applied for a three-years extension of the associate team.

6.1.3. Inverse magnetization issues from sparse spherical data

The team APICS is a partner of the ANR project MagLune concerning Lunar magnetism, associated to the Geophysics and Planetology Department of Cerege, CNRS, Aix-en-Provence (see Section 8.2.2). Measurements of the remanent magnetic field of the Moon let geoscientists think that the Moon used to have a magnetic dynamo for some time, but the exact process that triggered and fed this dynamo is not yet understood, much less why it stopped. In particular, the Moon is too small to have a convecting dynamo like the Earth has. The overall goal of the project is to devise models to explain how this dynamo phenomenon was possible on the Moon.

To this end, the geophysicists from Cerege will go to NASA to perform some measurements on samples brought back from the Moon by Apollo missions. The samples are kept inside bags with a protective atmosphere, and geophysicists are not allowed to open the bags, nor to take out the samples from NASA facilities. Therefore, measurements must be performed with some rudimentary instrument and our colleagues from Cerege designed a specific magnetometer. This device allows them to obtain measurements of the components of the magnetic field produced by the sample, at some discrete set of points located on disks belonging to three cylinders (see Figure 4).

This collaboration started this year and some preparatory work was necessary fix conventions used by our colleagues from Cerege in order to handle their measurements. During his Master 2 internship, Konstantinos Mavreas has developed a method based on rational approximation, using the same ideas as those underlying the FindSources3D tool (see Sections 3.4.2 and 6.1.1), for the case where the field produced by the sample can be well explained by a single magnetic dipole, whose position and moment are unknown. See his report ⁰. Konstantinos Mavreas is now engaged in a PhD within APICS and will extend these results to the case of several dipoles.

6.2. Matching problems and their applications

Participants: Laurent Baratchart, Martine Olivi, David Martinez Martinez, Fabien Seyfert.

This is collaborative work with Stéphane Bila (Xlim, Limoges, France), Yohann Sence (Xlim, Limoges, France), Thierry Monediere (Xlim, Limoges, France), Francois Torrès (Xlim, Limoges, France).

Filter synthesis is usually performed under the hypothesis that both ports of the filter are loaded on a constant resistive load (usually 50 Ohm). In complex systems, filters are however cascaded with other devices, and end up being loaded, at least at one port, on a non purely resistive frequency varying load. This is for example the case when synthesizing a multiplexer: each filter is here loaded at one of its ports on a common junction. Thus, the load varies with frequency by construction, and is not purely resistive either. Likewise, in an emitter-receiver, the antenna is followed by a filter. Whereas the antenna can usually be regarded as a resistive load at some frequencies, this is far from being true on the whole pass-band. A mismatch between the antenna and the

⁰<http://www-sop.inria.fr/apics/IMPINGE/>

⁰<http://www-sop.inria.fr/members/Konstantinos.Mavreas/main.pdf>

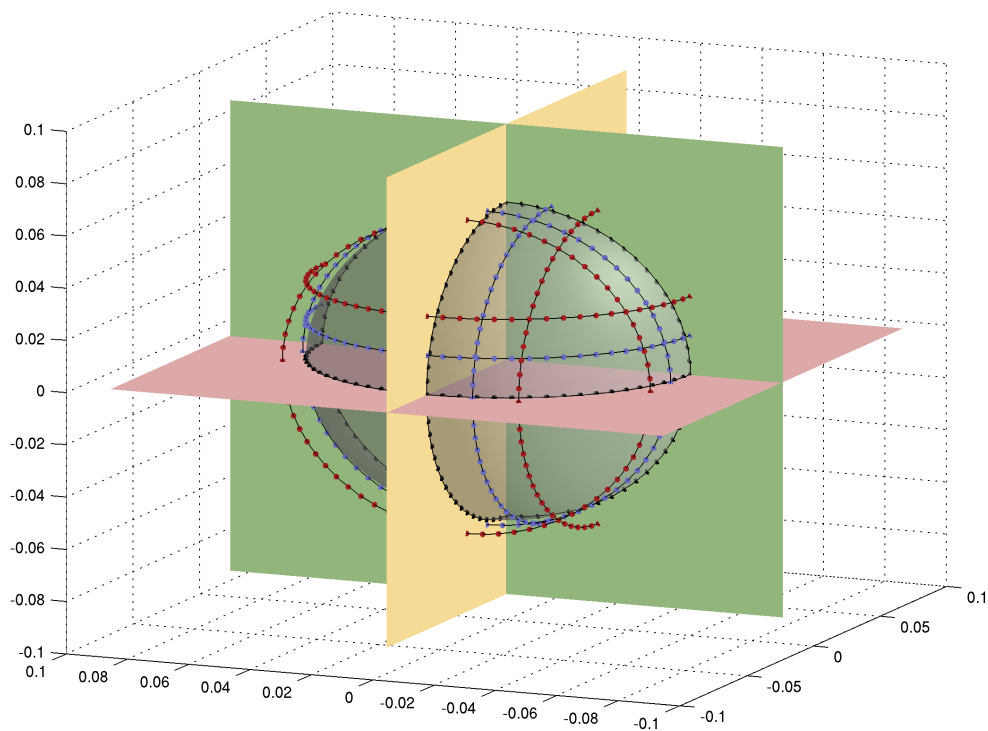


Figure 4. Typical measurements obtained with the instrument of Cerege. Discrete measurements of the field are performed on three cylinders. On each cylinder, the magnetic field \mathbf{B} is expressed as a component B_h co-linear with the axis of the cylinder, and a component \mathbf{B}_s parallel to a section of the cylinder. \mathbf{B}_s is itself decomposed as a tangential component B_τ and a normal component B_n , with respect to the circle given by the intersection of the cylinder with the corresponding section. At black points B_n is measured, at blue points B_h is measured, and at red points B_τ is measured.

filter, however, causes irremediable power losses, both in emission and transmission. Our goal is therefore to develop a method for filter synthesis that allows us to match varying loads on specific frequency bands, while enforcing some rejection properties away from the pass-band.

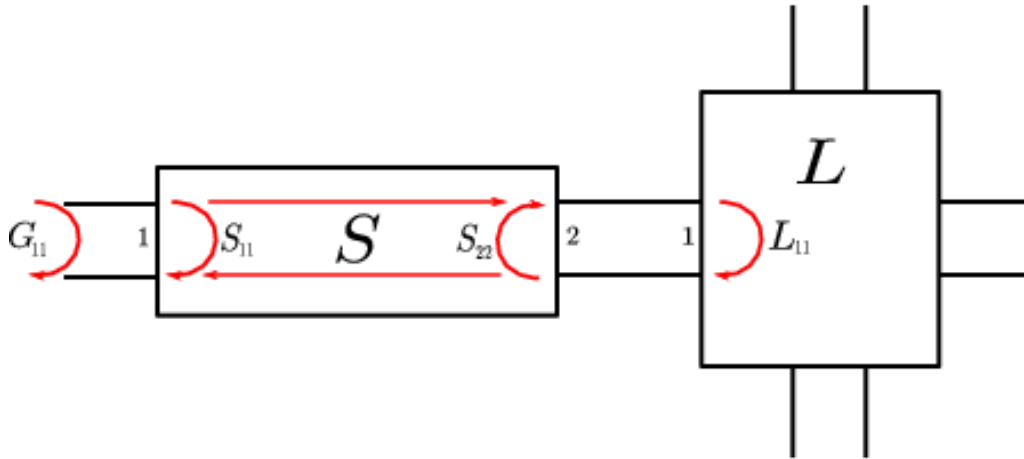


Figure 5. Filter plugged on a system with reflexion coefficient L_{11}

Figure 5 shows a filter with scattering matrix S , plugged at its right port on a frequency varying load with reflexion parameter $L_{1,1}$. If the filter is lossless, simple algebraic manipulations show that on the frequency axis the reflexion parameter satisfies:

$$|G_{1,1}| = \left| \frac{S_{2,2} - \overline{L_{1,1}}}{1 - S_{2,2}L_{1,1}} \right| = \delta(G_{1,1}, S_{2,2}).$$

The matching problem of minimizing $|G_{1,1}|$ amounts therefore to minimize the pseudo-hyperbolic distance δ between the filter's reflexion parameter $S_{2,2}$ and the load's reflexion $L_{1,1}$, on a given frequency band. On the contrary enforcing a rejection level on a stop band, amounts to maintaining the value of $\delta(L_{1,1}, S_{2,2})$ above a certain threshold on this frequency band. For a broad class of filters, namely those that can be modeled by a circuit of n coupled resonators, the scattering matrix S is a rational function of McMillan degree n in the frequency variable. The matching problem thus appears to be a rational approximation problem in the hyperbolic metric.

6.2.1. Approach based on interpolation

When the degree n of the rational function $S_{2,2}$ is fixed, the hyperbolic minimization problem is non-convex and led us to seek methods to derive good initial guesses for classical descent algorithms. To this effect, if $S_{2,2} = p/q$ where p, q are polynomials, we considered the following interpolation problem \mathcal{P} : given n frequency points $w_1 \cdots w_n$ and a transmission polynomial r , to find a monic polynomial p of degree n such that:

$$j = 1..n, \quad \frac{p}{q}(w_j) = \overline{L_{1,1}(w_j)}$$

where q is the unique monic Hurwitz polynomial of degree n satisfying the Feldtkeller equation

$$qq^* = pp^* + rr^*,$$

which accounts for the losslessness of the filter. The frequencies (w_k) are perfect matching points, as $\delta(S_{2,2}(w_k), L_{1,1}(w_k)) = 0$ holds, while the real zeros (x_k) of r are perfect rejection points (i.e. $\delta(S_{2,2}(x_k), L_{1,1}(x_k)) = 1$). The interpolation problem is therefore a point-wise version of our original matching-rejection problem. The monic restriction on p and q ensures the realisability of the filter in terms of coupled resonating circuits. If a perfect phase shifter is added in front of the filter, realized for example with a transmission line on a narrow frequency band, these monic restrictions can be dropped and an interpolation point w_{n+1} added, thereby yielding another interpolation problem $\widehat{\mathcal{P}}$. Our main result, states that \mathcal{P} as well as $\widehat{\mathcal{P}}$ admit a unique solution. Moreover the evaluation map defined by $\psi(p) = (p/q(x_1), \dots, p/q(x_n))$ is a homeomorphism from monic polynomials of degree n onto \mathbb{D}^n (\mathbb{D} the complex open disk), and ψ^{-1} is a diffeomorphism on an open, connected, dense set of \mathbb{D}^n . This last property has shown crucial for the design of an effective computational procedure based on continuation techniques. Current implementation of the latter tackles instances of \mathcal{P} or $\widehat{\mathcal{P}}$ for $n = 10$ in less than 0.1 *sec*, and allows for us a recursive use of this interpolation framework in multiplexer synthesis problems. We presented these techniques at the European Microwave Week 2015 in the workshop dedicated to "Recent Advances in the Synthesis of Microwave Filters and Multiplexers". The detailed mathematical proofs can be found in [21] and will be submitted shortly. On a related topic, namely the de-embedding of filters in multiplexers, our work has been published in [13].

6.2.2. Uniform matching and global optimality considerations

The previous interpolation procedure provides us with a matching/rejecting filtering characteristics at a discrete set of frequencies. This can serve as a starting point for heavier optimization procedures where the matching and rejection specifications are expressed uniformly over the bandwidth. Although the practical results thus obtained have shown to be quite convincing, we have no proof of their global optimality. This led us to seek alternative approaches able to assess, at least in simple cases, global optimality of the derived response. Following the approach of Fano and Youla, we considered the problem of a designing a 2×2 lossless frequency response, under the condition that a specified load can be "unchained" from one of its port. This classically amounts to set interpolation conditions on the response at the transmission zeros of the Darlington extension of the load. When the load admits a rational representation of degree 1, and if the transmission zeros of the overall system are fixed, then we were able to show that the uniform matching problem over an interval reduces to a convex minimization problem with convex constraints over the set of non-negative polynomials of given degree. In this case, which is already of some practical interest for antenna matching (antenna usually exhibit a single resonance in their matching band which is reasonably approximated at order 1), it is therefore possible to perform filter synthesis with a guarantee on the global optimality of the obtained characteristics. Procedures to derive the solution are currently being investigated, and lie at the heart of our contribution to the ANR-project Cocoram.

6.3. Unambiguous de-embedding of filters

Participants: Matthias Caenepeel, Martine Olivi, Fabien Seyfert.

This work was conducted in collaboration with Yves Rolain (VUB, Brussels, Belgium)

Coupling topologies that admit multiple realizations may lead to ambiguous de-embedding tuning procedures where distinct coupled resonator circuits are identified from the same measurements. This is for example the case of the well-known coupling topologies in triplets, quadruplets and extended boxed. If no additional measurements are performed on the DUT (device under tuning), the different solutions to the coupling matrix synthesis problem are undistinguishable, as they yield similar scattering responses. We therefore studied specific tuning strategies to discriminate among them. The later uses a sequence of measurements of the DUT, obtained after varying some discriminating tuning parameters of the filter and testing for coherence of the extracted circuits. This work was presented by Matthias Caenepeel at IMS 2015 in Phoenix [15] and at the

EuMC 2015 in Paris [16]. In a similar vein Matthias is currently developing techniques taking advantage of the differential information provided by EM solvers in order to compute the Jacobian matrix of the identified coupling matrix(ces) with respect to the geometrical parameters of the filter.

6.4. Orthogonal Polynomials

Participant: Laurent Baratchart.

We studied this year the asymptotic behavior of the orthonormal polynomials P_n with respect to a non-negative weight w on a simply connected planar domain Ω :

$$\int_{\Omega} P_n \bar{P}_k w \, dm = \delta_{n,k},$$

with $\delta_{n,k}$ the Kronecker symbol. We proved that if Ω has boundary $\partial\Omega$ of class $C^{1,\alpha}$, $\alpha > 0$, and if w converges in some appropriate sense to a boundary function $w_1 \in L^p(\partial\Omega)$ while not vanishing “too much” at the boundary, then

$$P_n(z) = \left(\frac{n+1}{\pi}\right)^{1/2} z^n S_{w_1}^-(\Phi(z)) \Phi^n(z) \Phi'(z) \{1 + o(1)\}$$

outside the convex hull of Ω , with Φ the conformal map from the complement of Ω onto the complement of the unit disk normalized so that $\Phi'(\infty) = \infty$, and $S_{w_1}^-$ the so-called exterior Szegő function of w_1 .

This generalizes considerably known asymptotics on analytic domains with Hölder smooth non vanishing weights [10]. The proof rests on some Hardy space theory, conformal mapping and $\bar{\partial}$ techniques. An exposition of the result was given at the conference *Orthogonal and Multiple Orthogonal Polynomials*, August 9-14 2015, Oaxaca (Mexico). An article is being written to report on this result.

6.5. Asymptotics of Rational Approximants

Participant: Laurent Baratchart.

This is joint work with M. Yattselev (IUPUI).

We studied best rational approximants in the *sup* norm to an analytic function f on compact set K of the analyticity domain Ω with connected complement. We showed that if the function can be continued analytically except over a set of logarithmic capacity zero comprising at most finitely many branchpoints, then the n -th root of the approximation error converges as n goes large to $e^{-2/C}$, with C the minimal Green capacity in $\mathbb{C} \setminus K$ of a compact set E outside of which f is single valued. Moreover, if $C > 0$, the normalized counting measure of the poles converges to the Green equilibrium distribution on E . We are currently considering the case of infinitely many branchpoints so as to get a somewhat final result on weak asymptotics in rational approximation to functions with polar singular set.

The proof rests on a blend of AAK-theory and potential theory.

ASPI Project-Team

5. New Results

5.1. Adaptive multilevel splitting

Participants: Frédéric Cérou, Arnaud Guyader.

We have show last year that an adaptive version of multilevel splitting for rare events is strongly consistent and that the estimates satisfy a CLT (central limit theorem), with the same asymptotic variance as the non-adaptive algorithm with the optimal choice of the parameters. This year we have generalized these results to include Markov kernels used to move the particles (or *shakers*) are of Metropolis–Hastings type. This is a non-trivial generalization to a very important case.

5.2. Adaptive multilevel splitting as a Fleming–Viot system

Participants: Frédéric Cérou, Arnaud Guyader.

This is a collaboration with Bernard Delyon (université de Rennes 1) and Mathias Rousset (EPI MATHERIALS, Inria Paris Rocquencourt).

By considering the adaptive multilevel splitting algorithm as a Fleming–Viot particle system for a stochastic wave, in the sense of [42], we have shown the mean square convergence using a general result [67] about the convergence of Fleming–Viot (Villemonais, 2013). We are currently working on the proof of a central limit theorem, but the proof is not yet complete. We have nevertheless identified the expression of the asymptotic variance.

5.3. Bias and variance reduction in rare event simulation

Participant: François Le Gland.

This is a collaboration with Damien Jacquemart (ONERA, Palaiseau) and Jérôme Morio (ONERA, Toulouse).

In [17], we highlight a bias induced by the discretization of the sampled Markov paths in the splitting algorithm, and we propose to correct this bias using a deformation of the intermediate regions, as proposed in [48]. Moreover, we propose two numerical methods to design intermediate regions in the splitting algorithm that minimise the variance. One is connected with a partial differential equation approach, the other one is based on the discretization of the state space of the process.

5.4. Simulation-based algorithms for the optimization of sensor deployment

Participant: François Le Gland.

This is a collaboration with Christian Musso (ONERA, Palaiseau) and with Sébastien Paris (LSIS, université du Sud Toulon Var).

The problem considered here can be described as follows: a limited number of sensors should be deployed by a carrier in a given area, and should be activated at a limited number of time instants within a given time period, so as to maximize the probability of detecting a target (present in the given area during the given time period). There is an information dissymmetry in the problem: if the target is sufficiently close to a sensor position when it is activated, then the target can learn about the presence and exact position of the sensor, and can temporarily modify its trajectory so as to escape away before it is detected. This is referred to as the target intelligence. Two different simulation-based algorithms have been designed in [23] to solve separately or jointly this optimization problem, with different and complementary features. One is fast, and sequential: it proceeds by running a population of targets and by dropping and activating a new sensor (or re-activating a sensor already available) where and when this action seems appropriate. The other is slow, iterative, and non-sequential: it proceeds by updating a population of deployment plans with guaranteed and increasing criterion value at each iteration, and for each given deployment plan, there is a population of targets running to evaluate the criterion. Finally, the two algorithms can cooperate in many different ways, to try and get the best of both approaches. A simple and efficient way is to use the deployment plans provided by the sequential algorithm as the initial population for the iterative algorithm.

5.5. Kalman Laplace filtering

Participant: François Le Gland.

This is a collaboration with Paul Bui Quang (CEA, Bruyères-le-Châtel) and Christian Musso (ONERA, Palaiseau).

We propose in [21] a new nonlinear Bayesian filtering algorithm where the prediction step is performed like in the extended Kalman filter, and the update step is done thanks to the Laplace method for integral approximation. This algorithm is called the Kalman Laplace filter (KLF). The KLF provides a closed-form non-Gaussian approximation of the posterior density. The hidden state is estimated by the maximum a posteriori. We describe a way to alleviate the computation cost of this maximization, when the likelihood is a function of a vector whose dimension is smaller than the state space dimension. The KLF is tested on three simulated nonlinear filtering problems: target tracking with angle measurements, population dynamics monitoring, motion reconstruction by neural decoding. It exhibits a good performance, especially when the observation noise is small.

5.6. Combining analog method and ensemble data assimilation

Participants: François Le Gland, Valérie Monbet, Chau Thi Tuyet Trang.

This is a collaboration with Pierre Ailliot (université de Bretagne Occidentale), Ronan Fablet and Pierre Tandéo (Télécom Bretagne), Anne Cuzol (université de Bretagne Sud) and Bernard Chapron (IFREMER, Brest).

Nowadays, ocean and atmosphere sciences face a deluge of data from spatial observations, in situ monitoring as well as numerical simulations. The availability of these different data sources offer new opportunities, still largely underexploited, to improve the understanding, modeling and reconstruction of geophysical dynamics. The classical way to reconstruct the space-time variations of a geophysical system from observations relies on data assimilation methods using multiple runs of the known dynamical model. This classical framework may have severe limitations including its computational cost, the lack of adequacy of the model with observed data, modeling uncertainties. In [24], we explore an alternative approach and develop a fully data-driven framework, which combines machine learning and statistical sampling to simulate the dynamics of complex system. As a proof concept, we address the assimilation of the chaotic Lorenz-63 model. We demonstrate that a nonparametric sampler from a catalog of historical datasets, namely a nearest neighbor or analog sampler, combined with a classical stochastic data assimilation scheme, the ensemble Kalman filter and smoother, reach state-of-the-art performances, without online evaluations of the physical model.

5.7. Markov-switching vector autoregressive models

Participant: Valérie Monbet.

This is a collaboration with Pierre Ailliot (université de Bretagne Occidentale), Julie Bessac (Argonne National Laboratory, Chicago) and Julien Cattiaux (Météo-France, Toulouse).

Multivariate time series are of interest in many fields including economics and environment. The most popular tools for studying multivariate time series are the vector autoregressive (VAR) models because of their simple specification and the existence of efficient methods to fit these models. However, the VAR models do not allow to describe time series mixing different dynamics. For instance, when meteorological variables are observed, the resulting time series exhibit an alternance of different temporal dynamics corresponding to weather regimes. The regime is often not observed directly and is thus introduced as a latent process in time series models in the spirit of hidden Markov models. Markov switching vector autoregressive (MSVAR) models have been introduced as a generalization of autoregressive models and hidden Markov models. They lead to flexible and interpretable models. In this multivariate context, several questions occur.

- The discrete hidden variable also called regime has to be correctly defined. Indeed the regime can be local (e.g. link to a subset of the variables) or global (e.g. the same for all the variables). It can also be observed and inferred a priori or hidden. In the second case, it has to be estimated at the same time as the model parameters.

The question of the definition of the regime is investigated in [26] for the specific problem of multi site wind modeling.

- Markov Switching VAR models (MSVAR) suffer of the same dimensionality problem as VAR models. For large (and even moderate) dimensions, the number of autoregressive coefficients in each regime can be prohibitively large which results in noisy estimates. When the variables are correlated, which is the standard situation in multivariate time series, over-learning is frequent. The estimated parameters contains spurious non-zero coefficients and are then difficult to interpret. The predictions associated to the model are usually unstable. Collinearity causes also ill-conditioning of the innovation covariance. In [29], we propose a likelihood penalization method with hard thresholding for MSVAR models leading to sparse MSVAR. Both autoregressive matrices and precision matrices are penalized using smoothly clipped absolute deviation (SCAD) penalties.

5.8. Dependent time changed processes

Participant: Valérie Monbet.

This is a collaboration with Pierre Ailliot (université de Bretagne Occidentale), Bernard Delyon (université de Rennes 1) and Marc Prevosto (IFREMER, Brest).

Many records in environmental sciences exhibit asymmetric trajectories and there is a need for simple and tractable models which can reproduce such feature. In [25] we explore an approach based on applying both a time change and a marginal transformation on Gaussian processes. The main originality of the proposed model is that the time change depends on the observed trajectory. We first show that the proposed model is stationary and ergodic and provide an explicit characterization of the stationary distribution. This result is then used to build both parametric and non-parametric estimate of the time change function whereas the estimation of the marginal transformation is based on up-crossings. Simulation results are provided to assess the quality of the estimates. The model is applied to wave data and it is shown that the fitted model is able to reproduce important statistics of the data such as its spectrum and marginal distribution which are important quantities for practical applications. An important benefit of the proposed model is its ability to reproduce the observed asymmetries between the crest and the troughs and between the front and the back of the waves by accelerating the chronometer in the crests and in the front of the waves.

5.9. An efficient algorithm for video super-resolution based on a sequential model

Participant: Patrick Héas.

This is a collaboration with Angélique Drémeau (ENSTA Bretagne, Brest) and Cédric Herzet (EPI FLUMINANCE, Inria Rennes–Bretagne Atlantique)

In the work [27], we propose a novel procedure for video super-resolution, that is the recovery of a sequence of high-resolution images from its low-resolution counterpart. Our approach is based on a *sequential* model (i.e. each high-resolution frame is supposed to be a displaced version of the preceding one) and considers the use of sparsity-enforcing priors. Both the recovery of the high-resolution images and the motion fields relating them is tackled. This leads to a large-dimensional, non-convex and non-smooth problem. We propose an algorithmic framework to address the latter. Our approach relies on fast gradient evaluation methods and modern optimization techniques for non-differentiable/non-convex problems. Unlike some other previous works, we show that there exists a provably-convergent method with a complexity linear in the problem dimensions. We assess the proposed optimization method on several video benchmarks and emphasize its good performance with respect to the state of the art.

5.10. Reduced-order modeling of hidden dynamics

Participant: Patrick Héas.

This is a collaboration with Cédric Herzet (EPI FLUMINANCE, Inria Rennes–Bretagne Atlantique).

The objective of the paper [28] is to investigate how noisy and incomplete observations can be integrated in the process of building a reduced-order model. This problematic arises in many scientific domains where there exists a need for accurate low-order descriptions of highly-complex phenomena, which can not be directly and/or deterministically observed. Within this context, the paper proposes a probabilistic framework for the construction of POD–Galerkin reduced-order models. Assuming a hidden Markov chain, the inference integrates the uncertainty of the hidden states relying on their posterior distribution. Simulations show the benefits obtained by exploiting the proposed framework.

BIPOP Project-Team

6. New Results

6.1. The contact complementarity problem, and Painlevé paradoxes

Participants: Bernard Brogliato, Florence Bertails-Descoubes, Alejandro Blumentals.

The contact linear complementarity problem is an set of equalities and complementarity conditions whose unknowns are the acceleration and the contact forces. It has been studied in a frictionless context with possibly singular mass matrix and redundant constraints, using results on well-posedness of variational inequalities obtained earlier by the authors [26]. This is also the topic of the first part of the Ph.D. thesis of Alejandro Blumentals where the frictional case is treated as a perturbation of the frictionless case [37]. The contact LCP is directly related to the so-called Painlevé's paradox of contact mechanics. In collaboration with C. Liu (Beijing university PKU) some results have been obtained from the analysis of a compliant model in the limit [34]. It shows on the classical sliding rod system that the inconsistent mode yield to instantaneous transition to a sticking mode. This is quite coherent with previous results obtained by Le xuan Anh in 1991 on the Painlevé-Klein system (bilateral constraints with Coulomb friction). With R. Kikuuwe from Kyushu University, we have also proposed a new formulation of the Baumgarte's stabilisation method, for unilateral constraints and Coulomb's friction, which sheds new light on Painlevé paradoxes as well [29].

6.2. Analysis of compliant nonlinear contact models

Participants: Bernard Brogliato, Guillaume James, Alexandre Vieira.

The master thesis of A. Vieira consisted of the study of suitable numerical method for compliant contact/impact models like the Simon-Hunt-Crossley and the Kuwabara-Kono models. These two models extend Hertz' contact by adding a dissipative force that takes the form of nonlinear viscous friction (nonlinear spring/dashpot). The fact that the Kuwabara-Kono dissipation is non-Lipschitz requires particular care.

6.3. Discrete-time sliding mode control

Participants: Vincent Acary, Bernard Brogliato, Olivier Huber.

This topic concerns the study of time-discretized sliding-mode controllers. Inspired by the discretization of nonsmooth mechanical systems, we propose implicit discretizations of discontinuous, set-valued controllers [3]. This is shown to result in preservation of essential properties like simplicity of the parameters tuning, suppression of numerical chattering, reachability of the sliding surface after a finite number of steps, and disturbance attenuation by a factor h or h^2 [22]. This work was part of the ANR project CHASLIM. Within the framework of CHASLIM we have performed many experimental validations on the electropneumatic setup of IRCCyN (Nantes), which nicely confirm our theoretical and numerical predictions: the implicit implementation of sliding mode control, drastically improves the input and output chattering behaviours, both for the classical order-one ECB-SMC and the twisting algorithms [33], [42], [27], [28]. In particular the high frequency bang-bang controllers which are observed with explicit discretizations, are completely suppressed. The implicit discretization has been applied to the classical equivalent-based-control SMC, and also to the twisting sliding-mode controller. The case of a nonlinear controller is studied in [35].

6.4. Lur'e set-valued dynamical systems: State observers

Participants: Bernard Brogliato, Christophe Prieur.

Lur'e systems are quite popular in Automatic Control since the fifties. Set-valued Lur'e systems possess a static feedback nonlinearity that is a multivalued function. We study in [53], [32] state observers for particular Lur'e systems which are Moreau's sweeping processes modelling Lagrange dynamics with frictionless unilateral constraints.

6.5. Measure Driven ODEs

Participants: Bernard Brogliato, Christophe Prieur.

Measure driven Ordinary differential equations are analyzed in [31] from the point of view of input-to-state stability (ISS). This relies on the solution concept introduced by Bressan and Rampazzo. Lyapunov-like functions are used to characterize the ISS. The link with impulsive ODEs and switching systems is made.

6.6. Numerical analysis of multibody mechanical systems with constraints

This scientific theme concerns the numerical analysis of mechanical systems with bilateral and unilateral constraints, with or without friction [2]. They form a particular class of dynamical systems whose simulation requires the development of specific simulators.

6.6.1. Numerical time–integration methods for event-detecting schemes.

Participants: Vincent Acary, Bernard Brogliato, Mounia Haddouni.

The CIFRE thesis of M. Haddouni concerns the numerical simulation of mechanical systems subject to holonomic bilateral constraints, unilateral constraints and impacts. This work is performed in collaboration with ANSYS and the main goal is to improve the numerical time–integration in the framework of event-detecting schemes. Between nonsmooth events, time integration amounts to numerically solving a differential algebraic equations (DAE) of index 3. We have compared dedicated solvers (Explicit RK schemes, Half-explicit schemes, generalizes α -schemes) that solve reduced index formulations of these systems. Since the drift of the constraints is crucial for the robustness of the simulation through the evaluation of the index sets of active contacts, we have proposed some recommendations on the use of the solvers of dedicated to index-2 DAE. A manuscript has been submitted to Multibody System Dynamics.

6.6.2. Multibody systems with clearances (dynamic backlash)

Participants: Vincent Acary, Bernard Brogliato, Narendra Akadkhar.

The PhD thesis of N. Akadkhar under contract with Schneider Electric concerns the numerical simulation of mechanical systems with unilateral constraints and friction, where the presence of clearances in imperfect joints plays a crucial role. A first work deals with four-bar planar mechanisms with clearances at the joints, which induce unilateral constraints and impacts, rendering the dynamics nonsmooth. The objective is to determine sets of parameters (clearance value, restitution coefficients, friction coefficients) such that the system's trajectories stay in a neighborhood of the ideal mechanism (*i.e.* without clearance) trajectories. The analysis is based on numerical simulations obtained with the projected Moreau-Jean time-stepping scheme. These results have been reported in [47]. It is planned to extend these simulations to frictional cases and to mechanisms of circuit breakers.

6.7. Nonlinear waves in granular chains

Participants: Guillaume James, Bernard Brogliato, Alexandre Vieira.

Granular chains made of aligned beads interacting by contact (e.g. Newton's cradle) are widely studied in the context of impact dynamics and acoustic metamaterials. While much effort has been devoted to the theoretical and experimental analysis of solitary waves in granular chains, there is now an increasing interest in the study of breathers (spatially localized oscillations) in granular systems. Due to their oscillatory nature and associated resonance phenomena, static or traveling breathers exhibit much more complex dynamical properties compared to solitary waves. Such properties have strong potential applications for the design of acoustic metamaterials allowing to efficiently damp or deviate shocks and vibrations. Our contribution to this field is twofold. In the work [52], the existence of static breathers is analyzed in granular metamaterials consisting of hollow beads with internal masses. Using multiple scale analysis and exploiting the unilateral character of Hertzian interactions, we show that long-lived breather solutions exist but time-periodic breathers do not (breather solutions actually disperse on long time scales). Moreover, in a collaboration with Y. Starosvetsky and D. Meimukhin (Technion), we numerically study the persistence of traveling breathers in

granular chains with local potentials under the effect of contact damping. Using a viscoelastic damping model (Hertz-Kuwabara-Kono model), we show that breathers can be generated by simple impacts in granular chains made from various materials (breathers propagate over a significant number of sites before being damped). The design of an experimental setup to test these theoretical predictions is underway. Another series of works concerns more specifically the modeling and numerical analysis of dissipative impacts : introduction of appropriate variables and simplifications for different models of contact damping (James, Brogliato), and comparative tests for various numerical discretizations of the Hunt-Crossley and Kuwabara-Kono models (Vieira, Brogliato, James).

6.8. Traveling pulses in the Burridge-Knopoff model

Participants: Guillaume James, Jose Eduardo Morales Morales, Arnaud Tonnelier.

The Burridge-Knopoff model describes the earthquake faulting process through the interaction of two plates modeled as a chain of blocks elastically coupled subject to a friction force. We study the existence of soliton-like solutions for the excitable Burridge-Knopoff model with different friction forces. We report for the first time the propagation of a one-pulse solitary wave where the position of the blocks remains unchanged after the passage of the wave. Extensive numerical simulations are done for different friction laws and a systematic investigation of the influence of the pulling velocity and the coupling constant is done. For a piecewise linear frictional law, we prove the existence of a traveling pulse in the weak coupling limit. A lower bound of the propagation speed is derived together with results on the shape of the traveling wave.

6.9. Propagation in space-discrete excitable systems

Participant: Arnaud Tonnelier.

We introduce a simplified model of excitable systems where the response of an isolated cell to an incoming signal is idealized by a fixed pulse-shape function. When the total activity of the cell reaches a given threshold a signal is sent to its N neighbors. We show that a chain of such excitable cells is able to propagate a set of simple traveling waves where the time interval between the firing of two successive cells remains constant. A comprehensive study is done for a transmission line with $N = 2$ and $N = 3$. It is shown that, depending on initial conditions, the network may propagate traveling waves with different velocities. Some necessary conditions for multistationarity are derived for an arbitrary N .

6.10. Inverse modeling with contact and friction

6.10.1. Inverse statics of plates and shells with frictional contact

Participants: Florence Bertails-Descoubes, Romain Casati, Gilles Daviet.

We pursued our work on the static inversion of thin elastic shells, in the presence of contact and friction with an external object. We have shown how to formulate draping as a local constrained minimization problem, and we have generalized the adjoint method to this constrained case. These new results are included in Romain Casati's PhD thesis, defended in June 2015, and will be part of a paper to be submitted in 2016.

6.11. Continuum modeling of granular materials

6.11.1. Continuum modeling of granular materials

Participants: Florence Bertails-Descoubes, Gilles Daviet.

We have proposed a new numerical framework for the continuous simulation of dilatable materials with pressure-dependent (Coulomb) yield stress, such as sand or cement. Relying upon convex optimization tools, we have shown that the continuous equations of motion coupled to the macroscopic nonsmooth Drucker-Prager rheology can be interpreted as the exact analogous of the solid frictional contact problem at the heart of Discrete Element Methods (DEM), extended to the tensorial space. Combined with a carefully chosen finite-element discretization, this new framework allowed us to avoid regularizing the continuum rheology while benefiting from the efficiency of nonsmooth optimization solvers, mainly leveraged by DEM methods so far. Our numerical results were successfully compared to analytic solutions on model problems, such as the silo discharge, and we retrieved qualitative flow features commonly observed in reported experiments of the literature. This work is currently under review at the Journal of Non Newtonian Fluid Mechanics, and a preliminary version is available as a research report [43]. Furthermore, we have recently extended the approach to account for flows with a varying density, leveraging the Material Point Method to discretize the Drucker Prager yield criterion without linearization. This work will be submitted to ACM SIGGRAPH in 2016.

6.12. Nonsmooth optimisation and applications

6.12.1. Semidefinite programming and combinatorial optimization

Participant: Jérôme Malick.

We have worked with Frederic Roupin (Prof. at Paris XIII) and Nathan Krislock (Assistant Prof. at North Illinois University, USA) on the use of semidefinite programming to solve combinatorial optimization problems to optimality. Nathan was the guest of the team during 2 months (June/July).

We have worked on a generic semidefinite-based solver for solve binary quadratic optimization problems. Using the generality of the bounds proposed in [54]. Our article is in revision in ACM Transaction of Mathematical Software. Our solver and our data sets are available online at <http://lipn.univ-paris13.fr/BiqCrunch/>.

Specializing the method of the k-cluster problem, we have proposed in [51] an algorithm able to solve exactly k-cluster instances of size 160. In practice, our method works particularly fine on the most difficult instances (with a large number of vertices, small density and small k).

6.12.2. Stochastic optimization for electricity production

Participant: Jérôme Malick.

Everyday, electricity generation companies submit a generation schedule to the grid operator for the coming day; computing an optimal schedule is called the unit-commitment problem. In collaboration with W. van Ackooij from EDF, we have proposed in [44] a two-stage formulation of unit-commitment to better include the impact of renewable energies. We present a primal-dual decomposition approach to tackle large-scale instances of these two-stage problems, wherein both the first and second stage problems are full unit-commitment problems. We provide an analysis of the theoretical properties of the algorithm, as well as computational experiments showing the interest of the approach for real-life large-scale unit-commitment instances.

6.13. Robotics

6.13.1. Mobile manipulation by humanoid robots

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Alexander Sherikov, Jory Lafaye, Camille Brasseur, Nestor Alonso Bohorquez Dorante.

This year's contributions to the field of mobile manipulation by humanoid robots have been three-fold: a lexicographic MPC approach to the decision of using optional contacts when necessary to maintain balance (and only when necessary), a robust MPC approach to online generation of dynamic walking motion on uneven ground such as stairs, and an analysis of the role of viability and capturability in collision prevention, using once again a lexicographic MPC approach.

6.13.2. Reactive trajectory generation

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Saed Al Homsy.

The goal of the collaboration with Adept Technologies is to generate time optimal trajectories in the presence of moving obstacles in real time. Three approaches with increasing computational complexity have been proposed and validated experimentally. The cheapest approach begins with a standard bang-bang control which is time-optimal in the absence of obstacles, and simply projected on dynamic limits imposed by collision avoidance. This leads to reasonable results where collisions are explicitly avoided, but time-optimality is lost in the process. A more complex approach introduces an MPC scheme minimizing a weighted L1-norm, which is tuned to generate a time-optimal behavior in the absence of obstacles. In the presence of obstacles, time-optimality is once again lost, however, results are much improved with respect to the previous approach. The final, and most complex approach, considers time-optimality as a lexicographic objective: a lexicographic MPC scheme is proposed, which achieves time-optimality in the presence of obstacles, with reasonable online computation time. This work has been submitted to ICRA 2016.

CAGIRE Team

7. New Results

7.1. DNS of a jet in crossflow

One main achievement of this year is to have done our first DNS computations at third order with the Aerosol software. Two configurations of jet in cross flow have been computed: one with a hole direction aligned with the main flow direction (Fig. 3 -left), and another one with a 90-degree jet skidding (Fig. 3 -right). The first case has been validated by using analytical models of jet trajectory, and has also been compared with experiments made with our experimental bench MAVERIC. The comparison of experiments and DNS showed a good agreement.

The DNS database includes:

- The instantaneous flow at the vertices of the mesh.
- the instantaneous flow at some probes.
- The mean flow.
- The value of the Reynolds stress tensor in all the degrees of freedom.

7.2. Extension of the momentum interpolation method to low Mach Riemann problems

In a previous study [9], the momentum interpolation (MI) method was considered as a guideline to develop a Godunov-like flux scheme called AUSM-IT and able to preserve the acoustic energy at the discrete level for a low-order finite volume approach. This year, the MI method has been successfully extended to the case of low Mach flows featuring discontinuities [8]. The undesirable dispersive effect directly connected to the upwinding of the MI formulation of the face velocity has been corrected (up to second-order errors) by using a central interpolation of momentum in the face velocity definition.

7.3. Main features of highly underexpanded jets

Despite the numerous studies dealing with underexpanded jets, many aspects of their structure were not clearly described, particularly when one seeks for quantitative predictions. Since such flow configuration may be of interest in case of the accidental boring of an aeronautical combustion chamber, an exhaustive review of the main experimental papers dealing with underexpanded jets has been carried out [5]. This study aimed at clarifying the characteristics which were well known, from those where there is clearly a lack of confidence. Curiously enough, such a work has never been done and no exhaustive review was available on such a topic.

7.4. Formulation of a reference EB-RSM model

The Elliptic Blending Reynolds Stress Model (EB-RSM), originally proposed by Manceau & Hanjalic in 2002, has been subject to various modifications by several authors during the last decade, mainly for numerical robustness reasons. We have revisited all these modifications from the theoretical standpoint and investigated in detail their influence on the reproduction of the physical mechanisms at the origin of the influence of the wall on turbulence. Theoretical arguments and comparison with DNS results led to the selection of a recommended formulation for the EB-RSM model [7].

7.5. Development of a new enrichment method

A complex issue in multi-scale simulations is the necessity to *enrich* the solution at the interface between a region described at coarse grain (e.g., using RANS) and a region described at fine grain (e.g., using LES). In order to rapidly generate realistic fluctuations at the beginning of the LES region, we have proposed [4] a method of volumic forcing, the so-called ALF (Anisotropic Linear Forcing). In an overlap region, a time-dependent volume force is introduced into the filtered equations of motion in order to amplify the turbulent fluctuations in order that the LES field satisfy the statistics of the RANS solution, a method that proved simple, efficient and computationally cheap.

7.6. A new criterion to analyse hybrid RANS/LES approaches

Most of the available hybrid RANS/LES methods are completely empirical or based on a formalism which is not applicable in practical application, due to a mismatch between the statistical average and the spatial filtering in inhomogeneous flows. The lack of clear formalism leads to limitations in terms of modeling of the unresolved turbulent motion. We have established a criterion [6] to assess the equivalence between hybrid RANS/LES methods, called *H-equivalence*, that makes it possible to view different hybrid methods as models for the same system of equations: as a consequence, empirical hybrid methods, such as the detached-eddy simulation (DES), can be interpreted as a model for the subfilter stress involved in the *temporally filtered* Navier-Stokes equations, which is an answer to the issue raised above about the formalism underlying such methods.

CARDAMOM Team

7. New Results

7.1. High order discretizations on unstructured meshes

Participants: Héloïse Beaugendre [Corresponding member], Cécile Dobrzynski, Léo Nouveau, Mario Ricchiuto, Quentin Viville.

Our work on high order unstructured discretizations this year has pursued three main avenues:

- We have extended the team's previous work on the consistent residual based approximation of viscous flow equations to the framework of Immersed Boundary Methods (IBM). This is an increasingly popular approach in Computational Fluid Dynamics as it simplifies the mesh generation problem. In our work, we consider a technique based on the addition of a penalty term to the Navier-Stokes equations to account for the wall boundary conditions. To adapt the residual distribution method to the IBM, we developed a new formulation based on a Strang splitting approach in time. This approach, couples in a fully consistent manner an implicit asymptotically exact integration procedure of the penalization ODE, with the explicit residual distribution discretization for the Navier-Stokes equations, based on the method proposed in (Ricchiuto and Abgrall, *J.Comput.Phys* 229, 2010). The ODE integrator provides an operator which is exact up to orders η^2 , with η the penalty parameter assuming values of the order of 10^{-10} . To preserve the accuracy of the spatial discretization in the Navier-Stokes step, we have introduced, in vicinity of the penalised region, a modification of the solution gradient reconstruction required for the evaluation of the viscous fluxes. We have shown formally and numerically that the approach proposed is second order accurate for smooth solutions, and shown its potential when combined with unstructured mesh adaptation strategies w.r.t. the (implicitly described) solid walls. This work has been accepted on *Comp.Meth.Appl.Mech.Eng.* ;
- Another research axis consists in proposing a novel approach that allows to use p-adaptation with continuous finite elements. Under certain conditions, primarily the use of a residual distribution scheme, it is possible to avoid the continuity constraint imposed to the approximate solution, while still retaining the advantages of a method using continuous finite elements. The theoretical material, the complete numerical method and practical results show as a proof of concept that p-adaptation is possible with continuous finite elements. Its extension to penalized Navier-Stokes equations are under progress ;
- We have continued the study of the properties of residual based methods in the time dependent case. We have been able to further characterize one of the variants of the approach proposed in (Ricchiuto and Abgrall, *J.Comput.Phys* 229, 2010) in terms of preservation of the positivity of the density showing this property in practical applications involving the shallow water equations [130]. The impact of the simplified construction leading to these schemes has also been investigated. In particular, we have shown that despite the additional complexity associated to the inversion of the mass matrix, non-linear methods providing monotone solution and yet featuring linear mass matrices can be constructed [142]. These methods, have been shown to have some potential w.r.t. fully diagonal approaches as those used in (Ricchiuto and Abgrall, *J.Comput.Phys* 229, 2010), in terms of error as function of CPU time : the non-diagonal schemes showing error reductions up to one order of magnitude. Current work is devoted to the use of other multistage (defect correction type) and multistep (extrapolated methods) techniques, comparing them to space time approaches.

7.2. Modelling of free surface flows

Participants: Luca Arpaia, Stevan Bellec, Mathieu Colin, Sebastien de Brye, Andrea Filippini, Maria Kazolea, Mario Ricchiuto [Corresponding member].

We have introduced a new systematic method to obtain discrete numerical models for incompressible free-surface flows. The method consists in first discretizing the Euler equations with respect to the horizontal variables, keeping the vertical z variable and time continuous. We have focused so far on (continuous) Galerkin and finite element discretizations in the horizontal. We then perform an asymptotic analysis on the resulting semi-discrete system. Our initial result, has led to a new discrete approximation, which we have shown to be consistent with the Boussinesq system known as Peregrine model. We have proven that the method obtained by means of this discrete asymptotic method, has phase and linear shoaling errors far lower than those obtained by discretizing the continuous model directly by means of the Galerkin method. Extensions to other weakly non-linear models have been obtained, and the study of fully nonlinear variants is under way.

We have also investigated the relations between some of the most common weakly nonlinear Boussinesq models. It is known since many years that, for given phase linear shoaling relations, two families of models exist depending on whether the dispersive terms are evaluated using derivatives of the speed, or of the flux (depth times speed). We have shown both analytically and numerically, that, independently on the phase and linear shoaling relations, these two families provide (qualitatively *and* quantitatively) only two distinct behaviours when approaching the nonlinear regime. Models based on velocity derivatives, provide taller more asymmetric waves, all models of the same family produce stunningly similar results, even when the linear relations differ considerably.

To extend our initial work on unstructured solvers for dispersive wave models to the fully nonlinear case we have proposed a new framework to approximate the so-called Green-Naghdi equations [99]. The method proposed, while remaining unsplit in time, is based on a separation of the elliptic and hyperbolic components of the equations. This separation is designed to recover the standard shallow water equations in the hyperbolic step, so that the method can be written as an *algebraic* correction to an existing shallow water code. In particular, in our approach we fix the method used for the elliptic component (a continuous Galerkin method), and couple it to different hyperbolic shallow water solvers. As long as the hyperbolic step is more than second order accurate, the approach proposed allows accuracies comparable to those of a fourth order finite difference method, with a natural potential for h and p - adaptation on unstructured grids. The two-dimensional extension is in the testing phase.

The tools developed have been also used intensively in funded research programs. Within the TANDEM project, several benchmarks relevant to tsunami modelling have been performed and several common publications with the project partners are in preparation. Independently on this activity, this year we used our codes to investigate two case studies. The first is the study of the wave conditions for the old Venetian harbour of Chania in Crete [109]. The study compares fully nonlinear-weakly dispersive COULWAVE code, developed at the University of South California, and TUCWave. The models are used to explore the appearance of resonance, eventually determining the resonant frequencies for the entire basin. The second study concerns the conditions for tidal bore formation in convergent alluvial estuaries [69]. A new set of dimensionless parameters has been introduced to describe the problem, and the code SLOWS has been used to explore the space of these parameters allowing to determine a critical curve allowing to characterize an estuary as "bore forming" or not. Surprising physical behaviours, in terms of dissipation and nonlinearity of the tides, have been highlighted. Part of this work has been accepted on *Estuarine, Coastal and Shelf Science*, with a manuscript on the numerical aspects in review on *Ocean Modelling*.

7.3. Wave energy conversion hydrodynamics

Participants: Umberto Bosi, Mario Ricchiuto [Corresponding member].

We have developed a prototype spectral element solver for a coupled set of differential equations modelling wave propagation (so-called outer domain), and the submerged flow under a floating body (inner domain). Both systems of equations are depth-averaged (Boussinesq type) systems involving some dispersive terms. They are further coupled to a force balance providing a (system of) ODE(s) for the floater. This model constitutes an intermediate fidelity approximation for the hydrodynamics of a wave energy converter. Differently from all industrial state of the art, it is a (fully) nonlinear model. However, its cost is extremely low when compared to full three-dimensional CFD analyses, due to the dimensional reduction brought from the depth averaged modelling. This year we have shown the potential of this Boussinesq-type model to predict the hydrodynamics

of a floater in a simplified case [97], [98] (journal version to appear on *J. Ocean Eng. and Marine Energy*). This result paves the way to the construction of new medium fidelity models to be used in the optimization of converters. This will be achieved in the framework of the MIDWEST project funded this year under the EU OCEANERanet call.

7.4. Two-phase flow numerical simulation with real-gas effects and occurrence of rarefaction shock waves

Participants: Maria Giovanna Rodio, Pietro Marco Congedo [Corresponding member].

We have studied the prediction in numerical simulation of turbulent cavitating flows, which could be strongly influenced by the presence of several empirical coefficients. The aim of this work is to explore the interaction between the cavitation model and turbulence in terms of uncertainty propagation through an unsteady numerical solver, for assessing the robustness and the accuracy of the physical models at different times. Furthermore, the influence of experimental data in the setting of some turbulence and cavitation model coefficients is investigated by means of a Bayesian approach. Finally, the interest is to provide some innovative insights for improving the understanding of these models for cavitating flows.

7.5. Formulation of stochastic methods for CFD

Participants: Maria Giovanna Rodio, Pietro Marco Congedo [Corresponding member].

We have worked on the extension of the Truncate and Encode (TE) approach, previously proposed by some of the authors (Abgrall et al. in *J Comput Phys* 257:19?56, 2014. doi:10.1016/j.jcp.2013.08.006), for taking into account uncertainty in partial differential equations (PDEs). Innovative ingredients are given by an algorithm permitting to recover the multiresolution representation without requiring the fully resolved solution, the possibility to treat a whatever form of pdf and the use of high-order (even non-linear, i.e. data-dependent) reconstruction in the stochastic space. Moreover, the spatial-TE method is introduced, which is a weakly intrusive scheme for uncertainty quantification (UQ), that couples the physical and stochastic spaces by minimizing the computational cost for PDEs. The proposed scheme is particularly attractive when treating moving discontinuities (such as shock waves in compressible flows), even if they appear during the simulations as it is common in unsteady aerodynamics applications. The proposed method is very flexible since it can easily coupled with different deterministic schemes, even with high-resolution features. Flexibility and performances of the present method are demonstrated on various numerical test cases (algebraic functions and ordinary differential equations), including partial differential equations, both linear and non-linear, in presence of randomness. The efficiency of the proposed strategy for solving stochastic linear advection and Burgers equation is shown by comparison with some classical techniques for UQ, namely Monte Carlo or the non-intrusive polynomial chaos methods.

A new scheme for the numerical approximation of a five-equation model taking into account Uncertainty Quantification (UQ) is also presented. In particular, the Discrete Equation Method (DEM) for the discretization of the five-equation model is modified for including a formulation based on the adaptive Semi-Intrusive (aSI) scheme, thus yielding a new intrusive scheme (sDEM) for simulating stochastic two-phase flows. Some reference test-cases are performed in order to demonstrate the convergence properties and the efficiency of the overall scheme. The propagation of initial conditions uncertainties is evaluated in terms of mean and variance of several thermodynamic properties of the two phases.

7.6. Sensitivity analysis, metamodelling, and and robust optimization

Participants: Kunkun Tang, Francesca Fusi, Pietro Marco Congedo [Corresponding member].

We have worked on two different formulations for sensitivity analysis. Moreover, we have proposed a new metamodelling technique and an innovative method for performing robust optimization.

Concerning sensitivity analysis, an anchored analysis of variance (ANOVA) method is proposed to decompose the statistical moments. Compared to the standard ANOVA with mutually orthogonal component functions, the anchored ANOVA, with an arbitrary choice of the anchor point, loses the orthogonality if employing the same measure. However, an advantage of the anchored ANOVA consists in the considerably reduced number of deterministic solver's computations, which renders the uncertainty quantification of real engineering problems much easier. Different from existing methods, the covariance decomposition of the output variance is used to take account of the interactions between non-orthogonal components, yielding an exact variance expansion and thus, with a suitable numerical integration method, provides a strategy that converges. This convergence is verified by studying academic tests. In particular, the sensitivity problem of existing methods to the choice of anchor point is analyzed via the Ishigami case, and we point out that covariance decomposition survives from this issue. Also, with a truncated anchored ANOVA expansion, numerical results prove that the proposed approach is less sensitive to the anchor point. The covariance-based sensitivity indices (SI) are also used, compared to the variance-based SI. Furthermore, we emphasize that the covariance decomposition can be generalized in a straightforward way to decompose higher-order moments. For academic problems, results show the method converges to exact solution regarding both the skewness and kurtosis. Finally, the proposed method is applied on a realistic case, that is, estimating the chemical reactions uncertainties in a hypersonic flow around a space vehicle during an atmospheric reentry.

A sensitivity analysis method is extended in order to compute third and fourth-order statistic moments, i.e. skewness and kurtosis, respectively. It is shown that this decomposition is correlated to a Polynomial Chaos (PC) expansion, permitting to easily compute each term. Then, new sensitivity indexes are proposed basing on the computation of skewness and kurtosis. PC-based numerical technique is used in order to compute the convergence of the sensitivity indexes according to the polynomial order by using the exact solution as the reference one. The interest of the proposed analysis is first depicted by considering several test-functions. In particular, a functional decomposition based on variance, skewness and kurtosis is computed, displaying how sensitivity indexes vary according to the order of the statistical moment. Then, the problem of how reducing the complexity of a stochastic problem is treated by proposing two strategies: i) the reduction of the number of dimensions, the reduction of the order of interaction. In both cases, the impact on the statistics of the reduced function is then assessed. Feasibility of the proposed analysis in a real-case is then demonstrated by presenting a stochastic study about the uncertainty propagation in a challenging engineering simulation: the numerical prediction of a turbine cascade in an Organic Rankine Cycles (ORCs), with the use of complex thermodynamic models and the presence of multiple sources of uncertainty. Basing on high-order statistics decomposition and physical remarks, it is shown how the analysis proposed in this work can help to drive the design process in a real-engineering problem.

For the metamodelling technique, a polynomial dimensional decomposition (PDD) method is proposed for the global sensitivity analysis and uncertainty quantification (UQ) of stochastic systems subject to a moderate to large number of input random variables. Due to the intimate structure between the PDD and the Analysis of Variance (ANOVA) approach, PDD is able to provide a simpler and more direct evaluation of the Sobol' sensitivity indices, when compared to the Polynomial Chaos expansion (PC). Unfortunately, the number of PDD terms grows exponentially with respect to the size of the input random vector, which makes the computational cost of standard methods unaffordable for real engineering applications. In order to address the problem of the curse of dimensionality, this work proposes essentially variance-based adaptive strategies aiming to build a cheap meta-model (i.e. surrogate model) by employing the sparse PDD approach with its coefficients computed by regression. Three levels of adaptivity are carried out in this paper: 1) the truncated dimensionality for ANOVA component functions, 2) the active dimension technique especially for second- and higher-order parameter interactions, and 3) the stepwise regression approach designed to retain only the most influential polynomials in the PDD expansion. During this adaptive procedure featuring stepwise regressions, the surrogate model representation keeps containing few terms, so that the cost to resolve repeatedly the linear systems of the least-squares regression problem is negligible. The size of the finally obtained sparse PDD representation is much smaller than the one of the full expansion, since only significant terms are eventually retained. Consequently, a much less number of calls to the deterministic model is required to compute the final PDD coefficients.

Concerning robust optimization, a strategy is developed to deal with the error affecting the objective functions in uncertainty-based optimization. We refer to the problems where the objective functions are the statistics of a quantity of interest computed by an uncertainty quantification technique that propagates some uncertainties of the input variables through the system under consideration. In real problems, the statistics are computed by a numerical method and therefore they are affected by a certain level of error, depending on the chosen accuracy. The errors on the objective function can be interpreted with the abstraction of a bounding box around the nominal estimation in the objective functions space. In addition, in some cases the uncertainty quantification methods providing the objective functions also supply the possibility of adaptive refinement to reduce the error bounding box. The novel method relies on the exchange of information between the outer loop based on the optimization algorithm and the inner uncertainty quantification loop. In particular, in the inner uncertainty quantification loop, a control is performed to decide whether a refinement of the bounding box for the current design is appropriate or not. In single-objective problems, the current bounding box is compared to the current optimal design. In multi-objective problems, the decision is based on the comparison of the error bounding box of the current design and the current Pareto front. With this strategy, fewer computations are made for clearly dominated solutions and an accurate estimate of the objective function is provided for the interesting, non-dominated solutions. The results presented in this work prove that the proposed method improves the efficiency of the global loop, while preserving the accuracy of the final Pareto front.

7.7. High order mesh generation and mesh adaptation

Participants: Luca Arpaia, Cécile Dobrzynski [Corresponding member], Ghina El Jannoun, Mario Ricchiuto.

This year several new algorithmic improvements have been obtained which will allow to enhance our meshing tools:

- We have enhanced our work on r -adaptation techniques for time dependent equations. These techniques are based on mesh deformations obtained by solving continuous differential equations for the local displacements. These equations are controlled by an error monitor. Several improvements have been made. We have studied in depth the formulation of the coupling of the mesh movement with the flow solver. We have found that for both finite volume and residual distribution methods, a coupling of mesh and solution evolution (by means of an ALE method) provides accuracy enhancements, and is to be preferred to a simpler adapt-project-evolve approach. The method has been fully tested in two space dimensions. The adaptation library has been extended to three dimensions, and benchmarking is under way. We have improved the definition of the error monitor, and we are now able to prescribe directly the local mesh size. For problems with source terms, and in particular problems admitting some important physical invariants as the shallow water equations, we have solved the conflict between the conservation of either mass or the invariant, allowing for the conservation of both quantities up to machine accuracy;
- We upgrade our technique for generating high order curved meshes: starting from a straight mesh with a curved boundary, a new smoothing and untangling approach is proposed to ensure a final valid mesh. The untangling algorithm is a hybrid technique that gathers a local mesh optimization applied on the surface of the domain and a linear elasticity analogy applied in its volume. On the one hand, the local topological optimization consists in simultaneously relocating the vertices and control points of a local patch around the invalid element in order to optimize the quality and validity of the elements inside the patch. The elements' validity problem is formulated as an unconstrained optimization problem using a log-barriers that is solved progressively using the conjugate gradient method. On the other hand, the linear elasticity analogy permits the propagation of the curvature to the volume of the domain hence untangling volume mesh elements.

7.8. Virtual self-healing composite for aeronautic propulsion

Participants: Mathieu Colin [Corresponding member], Xi Lin, Gregory Perrot, Mario Ricchiuto.

As a composite material having excellent properties, Ceramic-matrix composites (CMCs) comprise a ceramic matrix reinforced by a refractory fiber such as silicon carbide fiber. Due to the self-healing process (which consists in filling cracks produced by oxydation by an oxide formed in-situ), CMCs have extremely long lifetimes even under severe mechanical and chemical sollicitations. These time spans make most full-scale experimental investigations impractical : laboratory tests have necessarily to be replaced by predictions based on numerical models. Initial results have already been obtained by in the past with simplified crack averaged models based on simple potential approximations of the flow field of the oxide. In this direction, Xi Lin has developed new asymptotic models by creating a hierarchy inside two different families : the Saint-Venant equations and the thin film equations. The hierarchy is based on the use of several different boundary conditions. The main goal is to obtain more accurate hydrodynamic models accounting for surface tension and viscous effects which may be very important for the oxide evolution.

In parallel, we have made great progress in the coupling of the chemistry module with the structural mechanics solver of the LCTS laboratory in Bordeaux. The first fully coupled simulations of a fatigue test for a so-called mini-composite (on single fibre tow). The simulations have allowed to reproduce the gradual breaking mechanism typical of these materials, allowing to reproduce numerically the delayed rupture observed in practice.

7.9. Numerical simulation of the liquid ablation

Participants: Gérard Gallice, Luc Mieussens [Corresponding member], Simon Peluchon.

During the atmospheric re-entry phase, a space vehicle undergoes a heating due to the friction of the atmospheric gases. Conversion of kinetic energy to thermal energy leads to a sudden increase of the temperature of the solid boundary. This leads to a physical-chemical degradation of the thermal protective system, and to a boundary recession. For some materials, this recession occurs with a melting of the materials into a fluid phase. The numerical simulation of this phenomenon requires to take into account a two-phase flow with a compressible gaz (the air flow) and a weakly compressible liquid (the melted material). Numerically, this problem is strongly stiff.

We have proposed a splitting strategy to simulate compressible two-phase flows using the five equation model. The main idea of the splitting is to separate the acoustic and transport phenomena. The acoustic step is solved in Lagrangian coordinates by using a scheme based on an approximate Riemann solver. On the one hand, since the acoustic time step driven by the fast sound velocity is very restrictive, an implicit treatment of the Lagrangian step is performed. On the other hand, we use an explicit scheme for the transport step driven by the slow material waves. The global scheme resulting from this splitting operator strategy is conservative, positive, and preserves contact discontinuities. Numerical simulations of compressible diphasic flows are presented on 2d-structured grids. The implicit-explicit strategy allows large time steps, which do not depend on the fast acoustic waves.

7.10. Boundary conditions for the Navier-Stokes equations in the transitional regime

Participants: Céline Baranger, Pietro Marco Congedo, Giorgio Martalo, Julien Mathiaud, Luc Mieussens [Corresponding member].

In reentry flows at high altitude, the parietal fluxes along the boundary of a space vehicle are computed by solving the Boltzmann equation of the gas kinetic theory. It depends on the way particles are reflected by the solid wall. In this kind of applications, the reflection is usually supposed to be 80% diffuse (the particle are re-emitted in a random direction in thermal equilibrium with the wall), and 20% specular.

In lower altitude, it is possible to use the compressible Navier-Stokes equations, but the standard boundary conditions do not take into account the specular part. These equations are hence not very accurate in the transition regime (in which the Knudsen number is around 0.01).

By using an asymptotic boundary layer analysis, we have derived boundary conditions for the compressible Navier-Stokes equations that formally make the fluid model a second order approximation of the Boltzmann equation of the kinetic gas theory (with respect to the Knudsen number), and that can take into account the effect of the specular reflections. These boundary conditions include a slip velocity at the wall and a temperature jump, with some coefficients that depend on some auxiliary half-space problems. An existing numerical method has been extended to solve these problems and give numerical values for these coefficients. This makes our boundary conditions practically usable into any Navier-Stokes code.

COMMANDS Project-Team

7. New Results

7.1. Optimal control of ordinary differential equations

7.1.1. Periodic optimal controls for the Purcell microswimmer

Participant: Pierre Martinon.

We investigate in [31] some geometric and numerical aspects related to optimal control problems for the so-called Purcell Three-link swimmer, in which the cost to minimize represents the energy consumed by the swimmer. More precisely, we focus on the periodic aspect of optimal trajectories and controls. Linearizing the control system along a reference extremal, we estimate the conjugate points, which play a crucial role for the second order optimality conditions. With techniques imported by the sub-Riemannian geometry, we also show that the nilpotent approximation of the system provides a model which is integrable, obtaining explicit expressions in terms of elliptic functions. This approximation allows to compute optimal periodic controls for small deformations of the body. Numerical simulations are presented using Hampath and Bocop codes. A first paper was submitted in october 2015.

7.1.2. Study of optimal health insurance policies

Participant: Pierre Martinon.

In collaboration with the Economy department of Ecole Polytechnique, we analyze the design of an optimal medical insurance contract under ex post moral hazard, i.e., when illness severity cannot be observed by insurers and policyholders may exaggerate their health expenditures. This problem is reformulated in the optimal control framework, and we study the possible existence of deductibles or bunching phenomenons in optimal contracts. A paper will be submitted in early 2016.

7.2. Optimal control of partial differential equations

7.2.1. Local minimization algorithms for dynamic programming equations

Participant: Axel Kröner.

The numerical realization of the dynamic programming principle for continuous-time optimal control leads to nonlinear Hamilton-Jacobi-Bellman equations which require the minimization of a nonlinear mapping over the set of admissible controls. This minimization is often performed by comparison over a finite number of elements of the control set. In this paper we demonstrate the importance of an accurate realization of these minimization problems and propose algorithms by which this can be achieved effectively. The considered class of equations includes nonsmooth control problems with l_1 -penalization which lead to sparse controls. See the reprint [28].

7.2.2. Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids

Participant: Axel Kröner.

An approach to solve finite time horizon sub-optimal feedback control problems for partial differential equations is proposed by solving dynamic programming equations on adaptive sparse grids. The approach is illustrated for the wave equation. A semi-discrete optimal control problem is introduced and the feedback control is derived from the corresponding value function. The value function can be characterized as the solution of an evolutionary Hamilton-Jacobi Bellman (HJB) equation which is defined over a state space whose dimension is equal to the dimension of the underlying semi-discrete system. Besides a low dimensional semi-discretization it is important to solve the HJB equation efficiently to address the curse of dimensionality. We propose to apply a semi-Lagrangian scheme using spatially adaptive sparse grids. Sparse grids allow the discretization of the value functions in (higher) space dimensions since the curse of dimensionality of full grid methods arises to a much smaller extent. For additional efficiency an adaptive grid refinement procedure is explored. We present several numerical examples studying the effect the parameters characterizing the sparse grid have on the accuracy of the value function and the optimal trajectory. See the report [27].

7.2.3. Numerical approximation of level set power mean curvature flow

Participant: Axel Kröner.

In this paper we investigate the numerical approximation of a variant of the mean curvature flow. We consider the evolution of hypersurfaces with normal speed given by H^k , $k \geq 1$, where H denotes the mean curvature. We use a level set formulation of this flow and discretize the regularized level set equation with finite elements. In a previous paper we proved an a priori estimate for the approximation error between the finite element solution and the solution of the original level set equation. We obtained an upper bound for this error which is polynomial in the discretization parameter and the reciprocal regularization parameter. The aim of the present paper is the numerical study of the behavior of the evolution and the numerical verification of certain convergence rates. We restrict the consideration to the case that the level set function depends on two variables, i.e. the moving hypersurfaces are curves. Furthermore, we confirm for specific initial curves and different values of k that the flow improves the isoperimetrical deficit. See the report [29].

7.3. Finance and stochastic control

7.3.1. Second order Pontryagin's principle for stochastic control problems

Participant: Frédéric Bonnans.

In this Hal reprint [25], we discuss stochastic optimal control problems whose volatility does not depend on the control, and which have finitely many equality and inequality constraints on the expected value of functions of the final state, as well as control constraints. The main result is a proof of necessity of some second order optimality conditions involving Pontryagin multipliers.

7.3.2. On the convergence of the Sakawa-Shindo algorithm in stochastic control

Participant: Frédéric Bonnans.

In the accepted paper [32], we analyze an algorithm for solving stochastic control problems, based on Pontryagin's maximum principle, due to Sakawa and Shindo in the deterministic case and extended to the stochastic setting by Mazliak. We assume that either the volatility is an affine function of the state, or the dynamics are linear. We obtain a monotone decrease of the cost functions as well as, in the convex case, the fact that the sequence of controls is minimizing, and converges to an optimal solution if it is bounded. In a specific case we interpret the algorithm as the gradient plus projection method and obtain a linear convergence rate to the solution.

7.3.3. Optimal multiple stopping problems

Participant: Frédéric Bonnans.

In the paper [13] we extend some results by Carmona and Touzi [8], who studied an optimal multiple stopping time problem in a market where the price process is continuous. We generalize their results when the price process is allowed to jump. Also, we generalize the problem associated to the valuation of swing options to the context of jump diffusion processes. We relate our problem to a sequence of ordinary stopping time problems. We characterize the value function of each ordinary stopping time problem as the unique viscosity solution of the associated Hamilton–Jacobi–Bellman variational inequality. In the paper [14] we deal with numerical solutions to an optimal multiple stopping problem. The corresponding dynamic programming (DP) equation is a variational inequality satisfied by the value function in the viscosity sense. The convergence of the numerical scheme is shown by viscosity arguments. An optimal quantization method is used for computing the conditional expectations arising in the DP equation. Numerical results are presented for the price of swing option and the behavior of the value function.

7.4. Electricity production

7.4.1. Equilibria over energy markets

Participant: Benjamin Heymann.

Motivated by electricity markets we introduce in this paper a general network market model, in which agents are located on the nodes of a graph, a traded good can travel from one place to another through edges considering quadratic losses. An independent operator has to match locally production and demand at the lowest expense. As argued in our previous paper “Cost-minimizing regulations for a wholesale electricity market” this setting is relevant to describe some real electricity markets, pricing behavior and market power coming from the fact that generators can bid above their true value. In a general setting of many distributed generator agents connected by a transmission network, bidding piece-wise linear cost functions, we propose a pricing optimal mechanism model to reduce market power. Our main results are the expression of the optimal mechanism design, two algorithms for the allocation problem and market power estimations. To deduce these nice properties, we intensively use convex analysis and some monotone behaviors of the set-valued maps involved. Furthermore, these algorithms make it possible to numerically compute a Nash equilibrium for the procurement auction, which is important to compare the optimal mechanism and the standard auction setting. Finally, we also show some interesting examples. In the continuation of this work, we introduce a class of bidding games for which we prove the existence of a Nash equilibrium. We give a sufficient condition for uniqueness, propose a numerical scheme to compute the extreme Nash Equilibria and show that the equilibrium strategies are convex for a subclass of games. We apply this framework to electricity auctions.

7.4.2. Energy management for a micro-grid

Participants: Frédéric Bonnans, Benjamin Heymann, Pierre Martinon, Olivier Tissot.

We study in [33] the energy management problem for a microgrid including a diesel generator and a photovoltaic plant with a battery storage system. The objective is to minimize the total operational cost over a certain timeframe, primarily the diesel consumption, while satisfying a prescribed power load. After reformulation, the decision variables can be reduced to the charging /discharging power for the battery system. We take into account the switching cost for the diesel generator, the non-convex objective, and the long-term aging of the batteries. We solve this problem using a continuous optimal control framework, with both a direct transcription method (time discretization) and a Dynamic Programming method (Hamilton Jacobi Bellman). This project is a collaboration between team COMMANDS (Inria Saclay, France) and Centro de Energia (Universidad de Chile, Chile). Ongoing works include more refined battery aging models, and modeling the stochastic nature of the photovoltaic power and power load.

7.5. Energy management in transport

7.5.1. Energy management for an hybrid vehicle

Participants: Florine Bleuse, Frédéric Bonnans, Pierre Martinon.

In the framework of the PhD thesis of F.Bleuse, ‘Optimal control and robustness for rechargeable hybrid vehicles’. The study is focused on the so-called parallel architecture, with both the thermal and electric engines able to move the vehicle. The main axis is to optimize the use of the thermal engine. We started to develop a methodology with two time scales for solving the problem of computing a feedback control.

7.5.2. Collaboration with the startup Safety Line

Participants: Frédéric Bonnans, Pierre Martinon, Olivier Tissot.

We pursue our collaboration with Safety Line, using more refined atmospheric models (including for instance predicted wind data). Future works include high performance optimization for the cruise phase as well as analyzing the validity of the parameter estimation performed with the data from the flight recorders.

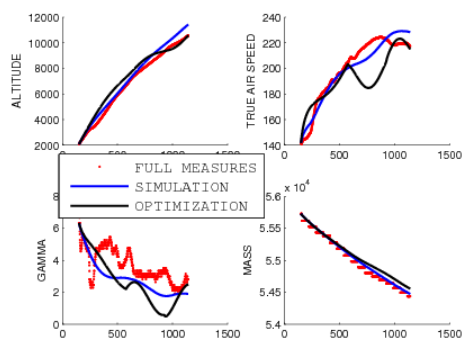


Figure 1. Parameter identification and trajectory optimization

CQFD Project-Team

7. New Results

7.1. Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies

The following result has been obtained by J. Anselmi (Inria CQFD), T. Nesti and B. Gaujal.

We consider a queueing system composed of a dispatcher that routes deterministically jobs to a set of non-observable queues working in parallel. In this setting, the fundamental problem is which policy should the dispatcher implement to minimize the stationary mean waiting time of the incoming jobs. We present a structural property that holds in the classic scaling of the system where the network demand (arrival rate of jobs) grows proportionally with the number of queues. Assume that each queue of type r is replicated k times and consider the set of policies that are periodic with period $k \sum_r p_r$ and such that exactly p_r jobs are sent in a period to each queue of type r . When $k \rightarrow \infty$, our main result shows that all the policies in this set are equivalent, in the sense that they yield the same mean stationary waiting time, and optimal, in the sense that no other policy having the same aggregate arrival rate to all queues of a given type can do better in minimizing the stationary mean waiting time. This property holds in a strong probabilistic sense. Furthermore, the limiting mean waiting time achieved by our policies is a convex function of the arrival rate in each queue, which facilitates the development of a further optimization aimed at solving the fundamental problem above for large systems.

7.2. Decentralized Proportional Load Balancing

The following result has been obtained by J. Anselmi (Inria CQFD), and N. Walton.

Load balancing is a powerful technique commonly used in communication and computer networks to improve system performance, robustness and fairness. In this paper, we consider a general model capturing the performance of communication and computer networks, and on top of it we propose a decentralized algorithm for balancing load among multiple network paths. The proposed algorithm is inspired by the modus operandi of the processor-sharing queue and on each network entry point operates as follows: every time a unit of load completes its service on a path, it increases by one unit the load of that path and decreases by one unit the load of a path selected at random with probability proportional to the amount of load on each of the available paths. We develop a dynamical system to argue that our load-balancer achieves a desirable network-wide utility optimization.

A paper has been accepted for publication in the SIAM Journal of Applied Mathematics.

7.3. Conditional quantile estimation through optimal quantization

The following result has been obtained by Isabelle Charlier (CQFD member), Davy Paindaveine, and Jérôme Saracco (CQFD member)

We use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d -dimensional vector of covariates X . First we focus on the population level and show how optimal quantization of X , which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X . We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed- N population counterpart. The results are illustrated on a numerical example. Dominance of our estimators over local constant/linear ones and nearest neighbor ones is demonstrated through extensive simulations in the companion paper Charlier et al. (2014).

7.4. A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes

The following result has been obtained by Oswaldo Costa and François Dufour (CQFD member).

This work deals with the constrained discounted control of piecewise deterministic Markov process (PDMPs) in general Borel spaces. The control variable acts on the jump rate and transition measure, and the goal is to minimize the total expected discounted cost, composed of positive running and boundary costs, while satisfying some constraints also in this form. The basic idea is, by using the special features of the PDMPs, to re-write the problem via an embedded discrete-time Markov chain associated to the PDMP and re-formulate the problem as an infinite dimensional linear programming (LP) problem, via the occupation measures associated to the discrete-time process. It is important to stress however that our new discrete-time problem is not in the same framework of a general constrained discrete-time Markov Decision Process and, due to that, some conditions are required to get the equivalence between the continuous-time problem and the LP formulation. We provide in the sequel sufficient conditions for the solvability of the associated LP problem. We provide some examples to illustrate the obtained results.

7.5. Impulsive control for continuous-time Markov decision processes

The following result has been obtained by Alexey Piunovskiy and François Dufour (CQFD member).

The objective of this work is to study continuous-time Markov decision processes on a general Borel state space with both impulsive and continuous controls for the infinite-time horizon discounted cost. The continuous-time controlled process is shown to be non explosive under appropriate hypotheses. The so-called Bellman equation associated to this control problem is studied. Sufficient conditions ensuring the existence and the uniqueness of a bounded measurable solution to this optimality equation are provided. Moreover, it is shown that the value function of the optimization problem under consideration satisfies this optimality equation. Sufficient conditions are also presented to ensure on one hand the existence of an optimal control strategy and on the other hand the existence of an ε -optimal control strategy. The decomposition of the state space in two disjoint subsets is exhibited where roughly speaking, one should apply a gradual action or an impulsive action correspondingly to get an optimal or ε -optimal strategy. An interesting consequence of our previous results is as follows: the set of strategies that allow interventions at time $t = 0$ and only immediately after natural jumps is a sufficient set for the control problem under consideration.

7.6. Impulsive control for continuous-time Markov decision processes: A Linear Programming Approach

The following result has been obtained by Alexey Piunovskiy and François Dufour (CQFD member).

The objective of this work is to investigate an optimization problem for continuous-time Markov decision processes with both impulsive and continuous controls. We consider the so-called constrained problem where the objective of the controller is to minimize a total expected discounted optimality criterion associated with a cost rate function while keeping other performance criteria of the same form, but associated with different cost rate functions, below some given bounds. Our model allows multiple impulses at the same time moment. The main objective of this work is to study the associated linear program defined on a space of measures including the occupation measures of the controlled process and to provide sufficient conditions to ensure the existence of an optimal control.

7.7. Conditions for the Solvability of the Linear Programming Formulation for Constrained Discounted Markov Decision Processes

The following result has been obtained by François Dufour (CQFD member) and T. Prieto-Rumeau.

This result concerns discrete-time constrained discounted Markov decision processes (MDP) with Borel state and action spaces, compact action sets, and lower semi-continuous cost functions. We introduce a set of hypotheses related to a positive weight function which allow us to consider cost functions that might not be bounded below by a constant, and which imply the solvability of the linear programming formulation of the constrained MDP. In particular, we establish the existence of a constrained optimal stationary policy. Our results are illustrated with an application to a fishery management problem.

7.8. Comparison of Kernel Density Estimators with Assumption on Number of Modes

The following result has been obtained by Gilles Durrieu, Raphaël Coudret and Jérôme Saracco (CQFD member).

A data-driven bandwidth choice for a kernel density estimator called critical bandwidth is investigated. This procedure allows the estimation to have as many modes as assumed for the density to estimate. Both Gaussian and uniform kernels are considered. For the Gaussian kernel, asymptotic results are given. For the uniform kernel, an argument against these properties is mentioned. These theoretical results are illustrated with a simulation study that compares the kernel estimators that rely on critical bandwidth with another one that uses a plug-in method to select its bandwidth. An estimator that consists in estimates of density contour clusters and takes assumptions on number of modes into account is also considered. Finally, the methodology is illustrated using environment monitoring data.

7.9. EEG classification for the detection of mental states

The following result has been obtained by Laurent Vezard, Pierrick Legrand (CQFD member), Marie Chavent (CQFD member), Frederique Faita-Ainseba and Trujillo Leonardo

The objective of the present work is to develop a method that is able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state of mind. For instance, pilots and medical staff are expected to be in a highly alert state and the proposed method could help to detect possible deviations from this expected state. This work poses a binary classification problem where the goal is to distinguish between a "relaxed" state and a baseline state ("normal") from the study of electroencephalographic signals (EEG) collected with a small number of electrodes. The EEG of 58 subjects in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select a subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the subjects. Different subset sizes were investigated and the best compromise between the number of selected electrodes and the quality of the solution was obtained by considering 9 electrodes. Even if the present approach is costly in computation time (GA search), it allows to construct a decision rule that provides an accurate and fast prediction of the alertness state of an unseen individual.

7.10. Modeling and optimization of a launcher integration process

The following result has been obtained by Christophe Nivot (CQFD member), Benoîte De Saporta, François Dufour (CQFD member), Jacques Béhar, Damien Bérard-Bergery and Charles Elegbede.

We deal with the modeling and the optimization of a launcher integration process. The subassemblies go through various types of operations which are split up into workshops. Their operating time is supposed random due to possible breakdowns or staff issues. Storage capacity of output products is limited and costly. Launches have to be performed according to a predetermined schedule, and lateness also costs money. The rate of production of the subassemblies must be decided every year. Therefore, the system can be modeled with a Markov decision process which is suitable for decision optimization and cost minimization. Indeed, one must find a balance between slow production (thus low storage levels and high probability to be late), and fast production (high storage levels but respected schedule).

We propose a model of this integration process based on Markov decision models. We present the simulation we have performed so far and discuss the difficulties of the optimization.

7.11. ClustGeo: Ascendant Hierarchical Clustering (AHC) with geographical constraints

The following result has been obtained by Marie Chavent (CQFD member), Vanessa Kuentz-Simonet, Amaury Labenne and Jerome Saracco (CQFD member).

Hierarchical Ascendant Clustering (HAC) is a well-known method of individual clustering. This method aims to bring together individuals who are similar regarding to variables which describe them. But when individuals are geographical units, the user may wish geographically close individuals to be put in same clusters and that, without too much deteriorating the quality of the partition. The proposed ClustGeo method allows geographical constraints of proximity to be taken into account within the HAC. For that purpose, a new Ward homogeneity criterion based on two different matrices of distances is proposed.

7.12. Approche bayésienne non paramétrique pour la factorisation de matrice binaire à faible rang avec loi de puissance

The following result has been obtained by Adrien Todeschini (CQFD member) and François Caron.

We introduce a low-rank Bayesian nonparametric (BNP) model for bipartite graphs. Recently, Caron (2012) proposed a BNP model where each node is given its own sociability parameter allowing to capture the power-law behavior of real world bipartite graphs. This model can be considered as a rank one nonnegative factorization of the adjacency matrix. Building on the compound random measures recently introduced by Griffin and Leisen (2014), we derive a rank p generalization of this model where each node is associated with a p -dimensional vector of sociability parameters accounting for several latent dimensions. While preserving the desired properties of interpretability, scalability and power-law behavior, our model is more flexible and provides better predictive performance as illustrated on several datasets.

7.13. Compétitions d'apprentissage automatique avec le package R rchallenge

The following result has been obtained by Adrien Todeschini (CQFD member) Robin Genuer.

In machine learning, empirical performance on real data are crucial in the success of a method. Recent years have seen the emergence of a large number of machine learning competitions. These challenges are motivated by industrial (Netflix prize) or academic (HiggsML challenge) applications and put in competition researchers and data scientists to obtain the best performance. We wanted to expose students to this reality by submitting a challenge in the context of the machine learning course. The leaderboard is displayed on an automatically updated web page allowing emulation among students. The history of the results also allows them to visualize their progress through the submissions. In addition, the challenge can continue outside of the supervised sessions promoting independence and exploration of new learning techniques and computer tools. The system we have implemented is available as an R package for reuse by other teachers. Building on the R Markdown and Dropbox tools, it requires no network configuration and can be deployed very easily on a personal computer.

7.14. Novelty Search

The following result has been obtained by Enrique Naredo, Leonardo Trujillo and Pierrick Legrand (CQFD member).

Novelty Search (NS) is a unique approach towards search and optimization, where an explicit objective function is replaced by a measure of solution novelty. However, NS has been mostly used in evolutionary robotics while its usefulness in classic machine learning problems has been unexplored. This work presents a NS-based Genetic Programming (GP) algorithm for supervised classification. Results show that NS can solve real-world classification tasks, validated on real-world benchmarks for binary and multiclass problems. These results are made possible by using a domain-specific behavior descriptor. Two new versions of the NS algorithm are proposed, Probabilistic NS (PNS) and a variant of Minimum Criterion NS (MCNS). The former models the behavior of each solution as a random vector and eliminates all of the original NS parameters while reducing the computational overhead of the NS algorithm. The latter uses a standard objective function to constrain and bias the search towards high performance solutions. The paper also discusses the effects of NS on GP search dynamics and code growth. Results show that NS can be used as a realistic alternative for supervised classification, and for binary problems the NS algorithm exhibits an implicit bloat control ability.

Keywords: Novelty Search, Behavior-based Search, Supervised Classification, Bloat

7.15. Classification of Epileptic states

The following result has been obtained by Emigdio Z. Flores, Leonardo Trujillo and Pierrick Legrand (CQFD member).

The neurological disorder known as Epilepsy is characterized by involuntary recurrent seizures that diminish a patient's quality of life. Automatic seizure detection can help improve a patient's interaction with her/his environment, and while many approaches have been proposed the problem is still not trivially solved. In this work, we present a novel methodology for feature extraction on EEG signals that allows us to perform a highly accurate classification of epileptic states. Specifically, Hölderian regularity and Matching Pursuit are used as the main feature extraction techniques, and are combined with basic statistics to construct the final feature sets. These sets are then delivered to a Random Forests classification algorithm. Furthermore, several versions of the basic problem are tested and statistically validated producing perfect accuracy in most problems and 92% accuracy on the most difficult case. A comparison with recent results in relevant literature using a well known database reveals that our proposal achieves state-of-the-art performance.

Keywords: Epilepsy detection, Hölderian regularity, Matching Pursuit, EEG Classification

7.16. Prediction of expected performance

The following result has been obtained by Yuliana Martinez, Leonardo Trujillo and Pierrick Legrand (CQFD member).

The study of problem difficulty is an open issue in Genetic Programming (GP). The goal of this work is to generate models that predict the expected performance of a GP-based classifier when it is applied to an unseen task. Classification problems are described using domain-specific features, some of which are proposed in this work, and these features are given as input to the predictive models. These models are referred to as predictors of expected performance (PEPs). We extend this approach by using an ensemble of specialized predictors (SPEPs), dividing classification problems into specified groups and choosing the corresponding SPEP. The proposed predictors are trained using 2D synthetic classification problems with balanced datasets. The models are then used to predict the performance of the GP classifier on unseen real-world datasets that are multidimensional and imbalanced. Moreover, as we know, this work is the first to provide a performance prediction of the GP classifier on test data, while previous works focused on predicting training performance. Accurate predictive models are generated by posing a symbolic regression task and solving it with GP. These results are achieved by using highly descriptive features and including a dimensionality reduction stage that simplifies the learning and testing process. The proposed approach could be extended to other classification algorithms and used as the basis of an expert system for algorithm selection.

7.17. Simulation of SPDEs for Excitable Media Using Finite Elements

The following result has been obtained by

M. Boulakia, A. Genadot (CQFD member) and M. Thioullien.

This result concerns the question of the discretization of Stochastic Partial Differential Equations (SPDE's) for excitable media. Working with SPDE's driven by colored noise, we consider a numerical scheme based on finite differences in time (Euler-Maruyama) and finite elements in space. Motivated by biological considerations, we study numerically the emergence of reentrant patterns in excitable systems such as the Barkley or Mitchell-Schaeffer models.

7.18. Conditional quantile estimation through optimal quantization: theoretical aspects

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, we use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d -dimensional vector of covariates X . First we focus on the population level and show how optimal quantization of X , which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X . We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed- N population counterpart. The results are illustrated on a numerical example.

7.19. Conditional quantile estimation based on optimal quantization: From theory to practice

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, small-sample properties of a nonparametric estimator of conditional quantiles based on optimal quantization, that was recently introduced (Charlier et al., JSPI, 2015), are investigated. More precisely, (i) the practical implementation of this estimator is discussed (by proposing in particular a method to properly select the corresponding smoothing parameter, namely the number of quantizers) and (ii) its finite-sample performances are compared to those of classical competitors. Monte Carlo studies reveal that the quantization-based estimator competes well in all cases and sometimes dominates its competitors, particularly when the regression function is quite complex. A real data set is also treated. While the main focus is on the case of a univariate covariate, simulations are also conducted in the bivariate case.

7.20. QuantifQuantile: An R Package for Performing Quantile Regression Through Optimal Quantization

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, we describe an R package, called QuantifQuantile, that allows to perform quantization-based quantile regression. In quantile regression, various quantiles of a response variable Y are modelled as functions of covariates (rather than its mean). An important application is the construction of reference curves/surfaces and conditional prediction intervals for Y . Recently, a nonparametric quantile regression method based on the concept of optimal quantization was proposed. This method competes very well with k -nearest neighbor, kernel, and spline methods. We describe also the various functions of the package and provide examples.

7.21. Numerical methods for simulation and optimization of piecewise deterministic Markov processes

This book is focused on theoretical and numerical aspects of simulation and optimization for piecewise deterministic Markov processes (PDMP's). PDMP's have been introduced in the literature by M. Davis as a general class of stochastic hybrid models. They form a family of Markov processes involving deterministic motion punctuated by random jumps. The motion of a PDMP includes both continuous and discrete variables. The continuous state variable represents the physical parameters of the system under consideration. The discrete mode characterizes the regimes of operation of the physical system and/or the environment. The process is defined through three local characteristics, namely the flow describing the deterministic trajectory between two consecutive jumps, the intensity function giving the jump rate and the Markov kernel specifying the post-jump location. A suitable choice of the state space and these local characteristics provides stochastic models covering a large number of problems such as engineering systems, operation research, economics, management science, biology, internet traffic, networks and reliability. The class of PDMP's is thus considered and recognized as a powerful modeling tool for complex systems.

However, surprisingly few works are devoted to the development of numerical methods for PDMP's to solve problems of practical importance such as evaluation and optimization of functionals of the process. The main objective of this book consists in presenting mathematical tools recently developed by the authors to address such problems. This book is not only focused on theoretical aspects such as proof of convergence of the approximation procedures but is also concerned with its applicability to practical problems. The approach we are proposing is general enough to be applied to several application domains. In particular, our results are illustrated by examples from the field of reliability.

Our approximation technique is based on the discretization using quantization of the underlying discrete-time Markov chain given by the post-jump locations and jump times of the PDMP. This strategy enables us to address a large class of numerical problems. In particular, in this book we focus, on the one hand, on the computation of expectation of functionals of PDMP's with applications to the evaluation of service times. On the other hand, we are interested in solving optimal control problems with applications to maintenance optimization.

DEFI Project-Team

6. New Results

6.1. Methods for inverse problems

6.1.1. Identifying defects in an unknown background using differential measurements

L. Audibert and H. Haddar

In the framework of the PhD thesis of Lorenzo Audibert we studied non destructive testing of concrete using ultrasonic waves, and more generally imaging in complex heterogeneous media. We assume that measurements are multistatic, which means that we record the scattered field on different points by using several sources. For this type of data we wish to build methods that are able to image the obstacle that created the scattered field. We use qualitative methods in this work, which only provide the support of the object independently from its physical property. The first part of this thesis consists of a theoretical analysis of the Linear Sampling Method. Such analysis is done in the framework of regularization theory, and our main contribution is to provide and analyze a regularization term that ensures good theoretical properties. Among those properties we were able to demonstrate that when the regularization parameter goes to zero, we actually construct a sequence of functions that strongly converges to the solution of the interior transmission problem. This behavior gives a central place to the interior transmission problem as it allows describing the asymptotic solution of our regularized problem. Using this characterization of our solution, we are able to give the optimal reconstruction we can get from our method. More importantly this description of the solution allows us to compare the solution coming from two different datasets. Based on the result of this comparison, we manage to produce an image of the connected component that contains the defect which appears between two measurement campaigns and this regardless of the medium. This method is well suited for the characteristics of the microstructure of concrete as shown on several numerical examples with realistic concrete-like microstructure. Finally, we extend our theoretical results to the case of limited aperture, anisotropic medium and elastic waves, which correspond to the real physics of the ultrasounds

6.1.2. Invisibility in scattering theory for small obstacles

L. Chesnel, X. Claeys and S.A. Nazarov

We are interested in a time harmonic acoustic problem in a waveguide containing flies. The flies are modelled by small sound soft obstacles. We explain how they should arrange to become invisible to an observer sending waves from $-\infty$ and measuring the resulting scattered field at the same position. We assume that the flies can control their position and/or their size. On the other hand, we show that any sound soft obstacle (non necessarily small) embedded in the waveguide always produces some non exponentially decaying scattered field at $+\infty$. As a consequence, the flies cannot be made completely invisible to an observer equipped with a measurement device located at $+\infty$.

6.1.3. New notion of regularization for Poisson data with an application to nanoparticle volume determination

F. Benvenuto, H. Haddar and B. Lantz

The aim of this work is to develop a fully automatic method for the reconstruction of the volume distribution of diluted polydisperse non-interacting nanoparticles with identical shapes from Small Angle X-ray Scattering measurements. The described method solves a maximum likelihood problem with a positivity constraint on the solution by means of an Expectation Maximization iterative scheme coupled with a robust stopping criterion. We prove that this is a regularization method according to an innovative notion of regularization specifically defined for inverse problems with Poisson data. Such a regularization, together with the positivity constraint results in high fidelity quantitative reconstructions of particle volume distributions making the method particularly effective in real applications. We test the performance of the method on synthetic data in the case of uni- and bi-modal particle volume distributions. Moreover, we show the reliability of the method on real data provided by a Xenocs device prototype.

6.1.4. A conformal mapping algorithm for the Bernoulli free boundary value problem

H. Haddar and R. Kress

We propose a new numerical method for the solution of Bernoulli's free boundary value problem for harmonic functions in a doubly connected domain D in R^2 where an unknown free boundary Γ_0 is determined by prescribed Cauchy data on Γ_0 in addition to a Dirichlet condition on the known boundary Γ_1 . Our main idea is to involve the conformal mapping method as proposed and analyzed by Akduman, Haddar and Kress for the solution of a related inverse boundary value problem. For this we interpret the free boundary Γ_0 as the unknown boundary in the inverse problem to construct Γ_0 from the Dirichlet condition on Γ_0 and Cauchy data on the known boundary Γ_1 . Our method for the Bernoulli problem iterates on the missing normal derivative on Γ_1 by alternating between the application of the conformal mapping method for the inverse problem and solving a mixed Dirichlet–Neumann boundary value problem in D . We present the mathematical foundations of our algorithm and prove a convergence result. Some numerical examples will serve as proof of concept of our approach.

6.1.5. Identification of small objects with near-field data in quasi-backscattering configurations

H. Haddar and M. Lakhali

We present a new sampling method for detecting targets (small inclusions or defects) immersed in a homogeneous medium in three-dimensional space, from measurements of acoustic scattered fields created by point source incident waves. We consider the harmonic regime and a data setting that corresponds with quasi-backscattering configuration: the data is collected by a set of receivers that are distributed on a segment centered at the source position and the device is swept along a path orthogonal to the receiver line. We assume that the aperture of the receivers is small compared with the distance to the targets. Considering the asymptotic form of the scattered field as the size of the targets goes to zero and the small aperture approximation, one is able to derive a special expression for the scattered field. In this expression a separation of the dependence of scattered field on the source location and the distance source-target is performed. This allows us to propose a sampling procedure that characterizes the targets location in terms of the range of a near-field operator constructed from available data. Our procedure is similar to the one proposed by Haddar-Rezac for far-field configurations. The reconstruction algorithm is based on the MUSIC (Multiple Signal Classification) algorithm.

6.2. Direct scattering problems

6.2.1. A numerical method to approximate black hole singularities in presence of metamaterials

L. Chesnel, A.-S. Bonnet-Ben Dhia, C. Carvalho and P. Ciarlet.

We investigate in a 2D setting the scattering of time-harmonic electromagnetic waves by a plasmonic device, represented as a non dissipative bounded and penetrable obstacle with a negative permittivity. Using the T-coercivity approach, we proved that the problem is well-posed in the classical frameworks if the negative permittivity does not lie in some critical interval whose definition depends on the shape of the device. When the latter has corners, for values inside the critical interval, unusual strong singularities for the electromagnetic field can appear. In that case, well-posedness is obtained by imposing a radiation condition at the corners to select the outgoing black-hole plasmonic wave, that is the one which carries energy towards the corners. We give a simple and systematic criterion to define what is the outgoing solution. We also propose an original numerical method based on the use of Perfectly Matched Layers at the corners. We emphasize that it is necessary to design an *ad hoc* technique because the field is too singular to be captured with standard finite element methods.

6.2.2. Boundary Integral Equations for the Transmission Eigenvalue Problem for Maxwell's Equations

Housseem Haddar, Shixu Meng and Fioralba Cakoni

We consider the transmission eigenvalue problem for Maxwell's equations corresponding to non-magnetic inhomogeneities with contrast in electric permittivity that changes sign inside its support. We formulate the transmission eigenvalue problem as an equivalent homogeneous system of boundary integral equation, and assuming that the contrast is constant near the boundary of the support of the inhomogeneity, we prove that the operator associated with this system is Fredholm of index zero and depends analytically on the wave number. Then we show the existence of wave numbers that are not transmission eigenvalues which by an application of the analytic Fredholm theory implies that the set of transmission eigenvalues is discrete with positive infinity as the only accumulation point.

6.2.3. A Volume integral method for solving scattering problems from locally perturbed periodic layers

Housseem Haddar and Thi Phong Nguyen

We investigate the scattering problem for the case of locally perturbed periodic layer in R^N ($N = 2, 3$). Using Floquet-Bloch transform in x_1 -direction we reformulate this scattering problem as an equivalent system of coupled volume integral equations. Using periodization in the x_2 -direction we apply a spectral method to discretize the problem and compute a numerical approximation of the solution. The convergence of this method is established and numerical validating results are conducted.

6.3. Shape and topology optimization

6.3.1. Deterministic approximation methods in shape optimization under random uncertainties

G. Allaire and C. Dapogny

This work is concerned with the treatment of uncertainties in shape optimization. We consider uncertainties in the loadings, the material properties, the geometry and the vibration frequency, both in the parametric and geometric optimization setting. We minimize objective functions which are mean values, variances or failure probabilities of standard cost functions under random uncertainties. By assuming that the uncertainties are small and generated by a finite number N of random variables, and using first- or second-order Taylor expansions, we propose a deterministic approach to optimize approximate objective functions. The computational cost is similar to that of a multiple load problems where the number of loads is N . We demonstrate the effectiveness of our approach on various parametric and geometric optimization problems in two space dimensions.

6.3.2. Molding direction constraints in structural optimization via a level-set method

G. Allaire, F. Jouve and G. Michailidis

In the framework of structural optimization via a level-set method, we develop an approach to handle the directional molding constraint for cast parts. A novel molding condition is formulated and a penalization method is used to enforce the constraint. A first advantage of our new approach is that it does not require to start from a feasible initialization, but it guarantees the convergence to a castable shape. A second advantage is that our approach can incorporate thickness constraints too. We do not address the optimization of the casting system, which is considered a priori defined. We show several 3d examples of compliance minimization in linearized elasticity under molding and minimal or maximal thickness constraints. We also compare our results with formulations already existing in the literature.

6.3.3. Identification of magnetic deposits in 2-D axisymmetric eddy current models via shape optimization

Zixian Jiang, Housseem Haddar, Armin Lechleiter and Mabrouka El-Guedri

The non-destructive control of steam generators is an essential task for the safe and failure-free operation of nuclear power plants. Due to magnetite particles in the cooling water of the plants, a frequent source for failures are magnetic deposits in the cooling loop of steam generators. From eddy current signals measured inside a U-tube in the steam generator, we propose and analyze a regularized shape optimization algorithm to identify magnetic deposits outside the U-tube with either known or unknown physical properties. Motivated by the cylindrical geometry of the U-tubes we assume an axisymmetric problem setting, reducing Maxwell's equations to a 2-D elliptic eddy current problem. The feasibility of the proposed algorithms is illustrated via numerical examples demonstrating in particular the stability of the method with respect to noise.

6.4. Asymptotic Analysis

6.4.1. *Ion transport through deformable porous media: derivation of the macroscopic equations using upscaling*

G. Allaire, O. Bernard, J.-F. Duf r che and A. Mikelic

We study the homogenization (or upscaling) of the transport of a multicomponent electrolyte in a dilute Newtonian solvent through a deformable porous medium. The pore scale interaction between the flow and the structure deformation is taken into account. After a careful adimensionalization process, we first consider so-called equilibrium solutions, in the absence of external forces, for which the velocity and diffusive fluxes vanish and the electrostatic potential is the solution of a Poisson-Boltzmann equation. When the motion is governed by a small static electric field and small hydrodynamic and elastic forces, we use O'Brien's argument to deduce a linearized model. Then we perform the homogenization of these linearized equations for a suitable choice of time scale. It turns out that the deformation of the porous medium is weakly coupled to the electrokinetics system in the sense that it does not influence electrokinetics although the latter one yields an osmotic pressure term in the mechanical equations. As a consequence, the effective tensor satisfies Onsager properties, namely is symmetric positive definite.

6.4.2. *On the asymptotic behaviour of the kernel of an adjoint convection-diffusion operator in a long cylinder*

G. Allaire and A. Piatnitski

This work studies the asymptotic behaviour of the principal eigenfunction of the adjoint Neumann problem for a convection diffusion operator defined in a long cylinder. The operator coefficients are 1-periodic in the longitudinal variable. Depending on the sign of the so-called longitudinal drift (a weighted average of the coefficients), we prove that this principal eigenfunction is equal to the product of a specified periodic function and of an exponential, up to the addition of fast decaying boundary layer terms.

6.4.3. *A comparison between two-scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures*

G. Allaire, M. Briane and M. Vanninathan

In this work we make a comparison between the two-scale asymptotic expansion method for periodic homogenization and the so-called Bloch wave method. It is well-known that the homogenized tensor coincides with the Hessian matrix of the first Bloch eigenvalue when the Bloch parameter vanishes. In the context of the two-scale asymptotic expansion method, there is the notion of high order homogenized equation where the homogenized equation can be improved by adding small additional higher order differential terms. The next non-zero high order term is a fourth-order term, accounting for dispersion effects. Surprisingly, this homogenized fourth-order tensor is not equal to the fourth-order tensor arising in the Taylor expansion of the first Bloch eigenvalue, which is often called Burnett tensor. Here, we establish an exact relation between the homogenized fourth-order tensor and the Burnett fourth-order tensor. It was proved by Conca et al. that the Burnett fourth-order tensor has a sign. For the special case of a simple laminate we prove that the homogenized fourth-order tensor may change sign. In the elliptic case we explain the difference between the homogenized and Burnett fourth-order tensors by a difference in the source term which features an additional corrector term.

Finally, for the wave equation, the two fourth-order tensors coincide again, so dispersion is unambiguously defined, and only the source terms differ as in the elliptic case.

6.4.4. Influence of the geometry on plasmonic waves

L. Chesnel, X. Claeys and S.A. Nazarov

In the modeling of plasmonic technologies in time harmonic regime, one is led to study the eigenvalue problem $-\operatorname{div}(\sigma \nabla u) = \lambda u$ (P), where σ is a physical coefficient positive in some region Ω_+ and negative in some other region Ω_- . We highlight an unusual instability phenomenon for the source term problem associated with (P): for certain configurations, when the interface between Ω_+ and Ω_- presents a rounded corner, the solution may depend critically on the value of the rounding parameter. We explain this property studying the eigenvalue problem (P). We provide an asymptotic expansion of the eigenvalues and prove error estimates. We establish an oscillatory behaviour of the eigenvalues as the rounding parameter of the corner tends to zero. These theoretical results are illustrated by numerical experiments.

6.4.5. Effective boundary conditions for thin periodic coatings Participants

Mathieu Chamaillard, Housseem Haddar and Patrick Joly

We study the derivation of asymptotic model (generalized impedance boundary conditions) for periodic coating in 3-D configurations. The definition of periodicity for 3D surfaces cannot be done in an intrinsic way in general. We propose a definition based on the use of local parametrisations of the surface. This parametrization-dependent definition is somehow inspired from practical considerations in the manufacturing of periodic coatings. The asymptotic of the problem is constructed for the scalar problem and also for the electromagnetic problem. Approximate models of order 1 and 2 are derived for the scalar problem and are validated numerically. In the electromagnetic case, only conditions of order 1 are exhibited in the general case.

6.5. Diffusion MRI

Jing-Rebecca Li, Housseem Haddar, Simona Schiavi, Khieu Van Nguyen, Gabrielle Fournet

Diffusion Magnetic Resonance Imaging (DMRI) is a promising tool to obtain useful information on microscopic structure and has been extensively applied to biological tissues.

We obtained the following results.

- We derived using homogenization techniques a model of the time-dependent “apparent diffusion coefficient” (ADC) that is valid at a wide range of diffusion times. The ADC is a very important experimental quantity measured by diffusion MRI in biological tissues. This work resulted in one submitted article to a mathematical journal and we are preparing an article for a physics journal.
- We analyzed a dMRI model called the Karger model that is valid at long diffusion times. This resulted in one submitted article to a mathematical journal.
- We acquired dMRI data of the nerve cells of the *Aplysia Californica* at the high field brain MRI center Neurospin. This data is useful because the nerve cells are bigger than mammal neurons and so it is easier to obtain segmented geometrical information about these cells for model validation.
- We participated in the data analysis and numerical simulation of a MR imaging method to measure blood flow in micro-vessels in the brain. This resulted in a submitted article to a MRI journal.

DISCO Project-Team

7. New Results

7.1. Algebraic Analysis Approach to Linear Functional Systems

Participants: Alban Quadrat [Disco], Rosane Ushirobira [Non-A].

7.1.1. Artstein's transformation of linear time-delay systems

Artstein's classical results show that a linear first-order differential time-delay system with delays in the input is equivalent to a linear first-order differential system without delays thanks to an invertible transform which includes integral and delay operators. Within a constructive algebraic approach, we show how Artstein's reduction can be found again and generalized as a particular isomorphism problem between the finitely presented modules defined by the two above linear systems over the ring of integro-differential time-delay operators. Moreover, we show that Artstein's reduction can be obtained in an automatic way by means of symbolic computation, and thus can be implemented in computer algebra systems.

7.1.2. Algebraic analysis for the Ore extension ring of differential time-varying delay operators

No algebraic (polynomial) approach seems to exist for the study of linear differential time-delay systems in the case of a (sufficiently regular) time-varying delay. Based on the concept of skew polynomial rings developed by Ore in the 30s, we construct the ring of differential time-delay operators as an Ore extension and to analyze its properties. A characterization of classical algebraic properties of this ring, such as noetherianity, its homological and Krull dimensions and the existence of Gröbner bases, are given in terms of the time-varying delay function. The algebraic analysis approach to linear systems theory allows us to study linear differential time-varying delay systems (e.g. existence of autonomous elements, controllability, parametrizability, flatness, behavioral approach) through methods coming from module theory, homological algebra and constructive algebra.

7.2. New Techniques for Robust Control of Linear Infinite-Dimensional Systems

Participants: Yacine Bouzidi [Disco], Petteri Laakkonen [Univ. Tampere], Adrien Poteaux [Lille 1], Alban Quadrat [Disco], Arnaud Quadrat [SAGEM], Guillaume Rance [SAGEM], Fabrice Rouillier [Ouragan].

7.2.1. Computer algebra methods for testing the structural stability of multidimensional systems

We present new computer algebra based methods for testing the structural stability of n -D discrete linear systems (with $n \geq 2$). More precisely, starting from the usual stability conditions which resumes to deciding if an hypersurface has points in the unit polydisc, we show that the problem is equivalent to deciding if an algebraic set has real points and use state-of-the-art algorithms for this purpose. Our strategy has been implemented in Maple and its relevance demonstrated through numerous experimentations.

Moreover, we also consider the specific case of two-dimensional systems and focus on the practical efficiency aspect. For such systems, the problem of testing the stability is reduced to that of deciding if a bivariate algebraic system with finitely many solutions has real ones. Our first contribution is an algorithm that answers this question while achieving practical efficiency. Our second contribution concerns the stability of two dimensional systems with parameters. More precisely, given a two-dimensional system depending on a set of parameters, we present a new algorithm that computes regions of the parameters space in which the considered system is structurally stable.

7.2.2. Computer algebra methods for the stability analysis of differential systems with commensurate time-delays

Within the frequency-domain approach, the asymptotic stability of linear differential systems with commensurate delays is ensured by the condition that all the roots of the corresponding quasipolynomial have negative real parts. A classical approach for checking this condition consists in computing the set of critical zeros of the quasipolynomial, i.e., the roots (and the corresponding delays) of the quasipolynomial that lie on the imaginary axis, and then analyzing the variation of these roots with respect to the variation of the delay. Based on solving algebraic systems techniques, we propose a certified and efficient symbolic-numeric algorithm for computing the set of critical roots of a quasipolynomial. Moreover, using recent algorithmic results developed by the computer algebra community, we present an efficient algorithm for the computation of Puiseux series at a critical zero which allows us to finely analyze the stability of the system with respect to the variation of the delay

7.2.3. A fractional ideal approach to the robust regulation problem

We show how fractional ideal techniques developed in [8] can be used to obtain a general formulation of the internal model principle for stabilizable infinite-dimensional SISO plants which do not necessarily admit coprime factorization. This result is then used to obtain necessary and sufficient conditions for the robust regulation problem. In particular, we find again all the standard results obtained in the literature.

7.2.4. Robust control as an application to the homological perturbation lemma:

Within the lattice approach to transfer matrices developed in [8], we have recently shown how standard results on robust control can be obtained in a unified way and generalized when interpreted as a particular case of the so-called Homological Perturbation Lemma. This lemma plays a significant role in algebraic topology, homological algebra, algebraic and differential geometry, computer algebra Our results show that it is also central to robust control theory for infinite-dimensional linear systems.

7.2.5. A symbolic-numeric method for the parametric H_∞ loop-shaping design problem

We develop a symbolic-numeric method for solving the H_∞ loop-shaping design problem for a low order single-input single-output system with parameters. Due to the system parameters, no purely numerical algorithm can indeed solve the problem. Using Gröbner basis techniques and the rational univariate representation of zero-dimensional algebraic varieties, we first give a parametrization of all the solutions of the two algebraic Riccati equations associated with the H_∞ control problem. Then, using results on the spectral factorization problem, a certified symbolic-numeric algorithm is obtained for the computations of the positive definite solutions of these two algebraic Riccati equations. Finally, we present a certified symbolic-numeric algorithm which solves the H_∞ loop-shaping design problem for the above class of systems.

7.3. Improved algorithm for computing separating linear forms for bivariate systems

Participants: Yacine Bouzidi [Disco], Sylvain Lazard [Vegas], Guillaume Moroz [Vegas], Marc Pouget [Vegas], Fabrice Rouillier [Ouragan].

We present new algorithms for computing linear separating forms, RUR decompositions and isolating boxes of the solutions. We show that these three algorithms have worst-case bit complexity $\tilde{O}_B(d^6 + d^5\tau)$, where \tilde{O} refers to the complexity where polylogarithmic factors are omitted and O_B refers to the bit complexity. We also present probabilistic Las-Vegas variants of our two first algorithms, which have expected bit complexity $\tilde{O}_B(d^5 + d^4\tau)$. A key ingredient of our proofs of complexity is an amortized analysis of the triangular decomposition algorithm via subresultants, which is of independent interest.

7.4. Stable H_∞ Controller for Infinite-dimensional systems

The controllers, besides the stabilization, are often designed to achieve some performance and robustness objectives by minimizing H_∞ norm of some cost functions. The resulting controller may be stable or unstable. The unstable controllers, however, are more sensitive to sensor/actuator faults, or nonlinearities. It is not an easy task to design a stable controller for systems having infinitely many zeros and poles in the right-half-plane. By using the similar idea in [88], stable H_∞ controller design method will be presented for a certain class of infinite-dimensional plants. The plants may have infinitely many unstable zeros, however, it is assumed that these zeros are *uniformly separated*. Under some certain assumptions, first, a sufficient condition will be presented to construct a real unit function, which satisfies certain interpolation conditions at the right-half-plane zeros of the plant and some H_∞ norm constraints. Then, by utilizing this result, stable H_∞ controller design method are presented.

7.5. Multiplicity and Stable Varieties of Time-delay Systems: A Missing Link

Multiple spectral values in dynamical systems are often at the origin of complex behaviors as well as unstable solutions. In this work, an unexpected property of multiple spectral values is emphasized. It has been shown that the variety corresponding to such a multiple root defines a stable variety for the steady state. Under mild assumptions, for the reduced examples we show that such a multiple spectral value is nothing else than the spectral abscissa.

7.6. Delay effect in chemical reactions

Belousov-Zhabotinskii (BZ) reaction, which is a very complicated reaction, has been widely studied in bio-science and chemistry, since its dynamic behaviour is similar to real biological oscillators [89]. For certain type of chemical reactions, the use of the law of mass-action kinetics may lead to some simple models expressed by ordinary differential equations. The main feature of BZ reaction, oscillatory behaviour, has been represented by a simple mechanism and the model of this mechanism can be described by ordinary differential equations. However, delayed mass-actions kinetics lead to more accurate models by conserving the simplicity and a relative reduced number of parameters. In [83], [95], [96], [82], some delay-differential models are proposed for Belousov-Zhabotinskii (BZ) reaction with a fewer of concentrations compared to the models obtained by ordinary differential equations. However, in most of these works, the delay, which occurs due to the required time to provide sufficient energy, has not been taken into account. Recently, we consider a more realistic Belousov-Zhabotinskii model, which includes two independent delays. The novelty of the proposed model with respect to the existing ones in the literature can be summarized as follows; one of these delays is introduced to reproduce qualitatively the behavior of the model proposed by [84] with a less number of concentrations as in [95]. Second, the remaining delay appears naturally since the reactants do not react suddenly in the chemical reactions, i.e. the delay stems from the needed time for the occurrence of reaction, called “delayed concentration”.

7.7. H_∞ -stability analysis of neutral systems with commensurate delays

Participants: Catherine Bonnet, Le Ha Vy Nguyen.

We have analyzed [32] the H_∞ -stability of neutral systems with commensurate delays and multiple chains of poles asymptotic to a same set of points on the imaginary axis. First, by approximation, the location of poles of large modulus is determined. This analysis requires to consider several subclasses of systems where poles of high modulus exhibit various patterns. Second, we derive necessary and sufficient conditions for H_∞ -stability which are easy to check as expressed in terms of the degrees of the polynomials involved in the numerator and denominator of the transfer function.

7.8. H_∞ -Stabilization of neutral delay systems

Participants: Catherine Bonnet, Yutaka Yamamoto [Kyoto University].

We have considered two particular neutral delay systems with one delay having a chain of poles clustering the imaginary axis from left or right. For these systems the existence of coprime factorizations have been investigated. The extension to more general systems is still in progress.

7.9. Interval Observer of a new type

Participants: Frederic Mazenc [Disco], Emilia Fridman [Tel-Aviv University].

In [19], we addressed the fundamental problem of constructing for nonlinear systems observers that converge in finite time and, at the same time, provided with upper and lower bounds for the solutions when disturbances are present. This new technique of estimation relies on the use of past values of the output, as done to construct some already known observers which converge in finite time, and on a recent technical result pertaining to the theory of the monotone systems. The result applies to systems with additive disturbances and disturbances in the output. The nonlinear terms are not supposed to be globally Lipschitz, but it is requested that they depend only on the input and output variables. The fundamental advantage over classical interval observer techniques is that no information on the initial conditions of the solutions of the studied system are needed.

7.10. Trajectory based approach

Participants: Frederic Mazenc [Disco], Silviu Niculescu [Disco], Michael Malisoff [LSU].

In the work [22], we provided a new stability analysis technique, which is based on the study of the behaviors of the solutions over any interval $[t - T, T]$, where t represents the time and T is an appropriately chosen constant. Thus trajectory-based approach is completely new in the sense that it neither rely on Lyapunov functions nor on the small gain theorem. One of its most striking feature is that it applies to a broad number of systems (systems with delay, continuous/discrete systems, ODE coupled with difference equations).

In [30] and [55] we provided several significant applications of the main result of [22]. In [30] in two results, we use a Lyapunov function for a corresponding undelayed system to provide a new method to prove stability of linear continuous-time time-varying systems with bounded time-varying delays. We allow uncertainties in the coefficient matrices of the systems. Our main results use upper bounds on an integral average involving the delay. The results establish input-to-state stability with respect to disturbances. We also provided in [55] a novel reduction model approach that ensures global exponential stabilization of linear systems with a time-varying pointwise delay in the input, which allows the delay to be discontinuous and uncertain. Finally, we provided an alternative to the reduction model method, based on a different dynamic extension.

7.11. Positive Systems Approach

Participants: Frederic Mazenc [Disco], Michael Malisoff [LSU].

We presented new methods to prove stability of time-varying systems with delays by taking advantage of the theory of the positive and cooperative systems [24], [23]. We used linear time-varying Lyapunov functionals, operators with integral terms, and positive systems, and we provided robustness of the stability with respect to multiplicative uncertainty in the vector fields. We allowed cases where the delay may be unknown, and where the vector fields defining the systems are not necessarily bounded. The results apply to neutral time-varying systems but are very distinct from those of the paper [69].

7.12. Attitude dynamics, control and observation

Participants: Frederic Mazenc [Disco], Maruthi Akella [Univ of Texas], Divya Thakur [Univ of Texas], Sunpil Yang [Univ of Texas].

We addressed several problems pertaining to the control of fully actuated rigid-body attitude dynamics. The fundamental tool we used is the adaptation of the so called strictification approach to the features of the attitude dynamics (see [3] for an introduction to the 'strictification' paradigm). In particular

1) The contribution [31] output feedback stabilization of fully actuated rigid-body attitude dynamics in the presence of unknown point-wise time-delay in the input torque. Specifically, rate-gyros are unavailable here and only the attitude state represented by the unit quaternion is assumed to be measured. It is worth mentioning that the presence of unknown time-delay in the measured variables, imposes formidable technical challenges for the output-feedback attitude stabilization problem on hand. One of the central difficulties stems from the availability of only a weak Lyapunov-like function for the passivity based dynamic output feedback controller in the absence of delay. This obstacle is circumvented in this contribution by a novel process of partially strictifying the underlying weak Lyapunov-like function.

2) In [57], we considered stabilization of fully actuated rigid-body attitude dynamics in the absence of angular velocity measurements and presents new robustness results to bounded unknown external disturbance torques. In particular, it is assumed that only body orientation is measured in the form of a unit-quaternion signal. It is well known that the passivity properties of the dynamics allows design of velocity-free controllers using a first-order stable filter driven by measured states. When external disturbance torques are taken into account, however, the robustness properties of these passivity-based output feedback controllers cannot be readily analyzed because the Lyapunov-like function from which the controller is derived has a time-derivative that is only negative semidefinite, and therefore non-strict. This obstacle is circumvented through a new partial-strictification approach which ultimately allows the characterization of robustness properties for this closed-loop system.

3) In [18], we proposed a smooth angular velocity observer for the attitude tracking control of a rigid body in the absence of angular velocity measurements. The observer design ensures asymptotic convergence of angular velocity state estimation errors irrespective of the control torque or the initial attitude state of the spacecraft. Unlike existing rate observer formulations that attain estimation error convergence by imposing certain switching conditions or hybrid-logic, the proposed observer has a smooth structure that ensures continuity of all estimated states. Lyapunov strictification is again the key technical result making us to establish the results.

7.13. Introduction of artificial delays for control and observation

Participants: Frederic Mazenc [Disco], Silviu Niculescu [Disco], Michael Malisoff [LSU], Nikolaos Bekiaris-Liberis [Tech. Univ. of Crete].

It is well-known that, in some cases, control or observation problems for systems with or without inherent delay can be solved by artificially introducing delays. We obtained in this field of research two distinct new results.

1) In [56] and [54], we have considered a family of linear time-varying systems with an input, an output and delays in the input. We have shown that, under classical stabilizability and detectability assumptions, all the systems of this family can be exponentially stabilized through a time-varying feedback depending on past values of the output and the input and this without the use of observers or dynamic extensions. Hence, the simplicity of the design and the determination of the value of the solutions in finite time are the main features of the new approach.

2) In [54], we provided a new backstepping result for time-varying systems with input delays. The novelty of the contribution is in the bounds on the controls, and the facts that (i) one does not need to compute any Lie derivatives to apply our controls, (ii) the controls have no distributed terms, and (iii) no differentiability conditions on the available controls for the subsystems are needed. The result is obtained by the introduction of constant pointwise delay in the input. Thus this result is significantly different for backstepping results for systems with delay in the input as presented for instance in [70].

7.14. Continuous-discrete observers

Participants: Frederic Mazenc [Disco], Emilia Fridman [Tel-Aviv University], Michael Malisoff [LSU], Vincent Andrieu [LAGEP].

We solved several problems of observer design pertaining to the fundamental and difficult case where the measurements are available at discrete instants only.

1) We considered the problem of stabilizing a linear continuous-time system with discrete-time measurements and a sampled input with a pointwise constant delay [20]. In a first part, we designed a continuous-discrete observer which converges when the maximum time interval between two consecutive measurements is sufficiently small. In a second part, we constructed a dynamic output feedback by using a technique which is strongly reminiscent of the reduction model approach. It stabilizes the system when the maximal time between two consecutive sampling instants is sufficiently small. No limitation on the size of the delay is imposed and an ISS property with respect to additive disturbances is established.

2) The problem of designing continuous-discrete observers for a large class of continuous time nonlinear time-varying systems with discrete time measurements has been addressed in several contributions: [12], [17] and [53]. Some technical obstacle encountered in [12] were overcome in [17] by using the notion of cooperative systems, which led to results consisting in explicit expressions of the largest sampling interval under which the observers converge to the solutions of the original system

7.15. Sensor faults: detectability and distributed Fault Diagnosis. A positive invariance based approach.

Participants: Sorin Olaru [correspondent], Vasso Reppas [L2S], Abid Kodakkadan [L2S], Marios Polycarpou [University of Cyprus].

The paper [62] introduces the performance analysis of local monitoring modules of a distributed diagnosis scheme tailored to detect multiple sensor faults in a class of nonlinear systems. The local modules monitor the healthy operation of subsets of sensors (local sensor sets). Every module is designed to detect the occurrence of faults in the local sensor sets when some analytical redundancy relations (ARRs) are violated. The set of ARR is formulated using structured residuals and adaptive thresholds based on a nonlinear observer. In order to characterize the sensitivity of every monitoring module to local sensor faults, we obtain structural fault detectability conditions based on adaptive thresholds, and strong fault detectability conditions based on ultimate robust positively invariant sets. These conditions correspond to explicit relationships between the local sensor faults, the worst-case bounds on modeling uncertainties and the design parameters of the local monitoring module.

In a recent paper [44], we considered the abnormal functioning of sensors (measurement channels) deployed for monitoring and control of discrete linear time invariant systems affected by additive uncertainties. The main objective was to analyze the sensor fault detectability via a robust positive invariance based technique. The analysis relies on the categorization of detectable faults and leads to certain conditions for guaranteed nondetectability, guaranteed detectability and implicit detectability.

As a support to this line of research, in the paper [73] we presented a methodology for computing robust positively invariant sets for linear, discrete time-invariant systems that are affected by additive disturbances, with the particularity that these disturbances are subject to state-dependent bounds. The proposed methodology requires less restrictive assumptions compared to similar established techniques, while it provides the framework for determining the state-dependent (parameterized) ultimate bounds for several classes of disturbances. The added value of the proposed approach is illustrated by an optimization-based problem for detecting the mode of functioning of a switching system.

7.16. Fault tolerant control design for multi-sensor networked control systems

Participants: Sorin Olaru [correspondent], Nikola Stankovic [L2S], Silviu Niculescu [L2S], Florin Stoican [Univ. Politehnica, Bucharest].

In the paper [27], we consider a multi-sensor networked control configuration with linear plant which is affected by a bounded additive disturbance. Shared network is used for the communication between sensors and controller. It is assumed that the sensors are prone to abrupt faults, while the controller's input may be updated with a varying time-delay. In order to identify and isolate the sensor(s) providing faulty information, we equip the controller with a set-based detection and isolation routine. Furthermore, in the case when the network induces time-delays, control is performed based on the knowledge we have on the mathematical model of the plant. In the presence of model inaccuracies or disturbance, such a control action may not guarantee satisfying performance of the system. Therefore, a stabilising controller with delay compensation has been designed.

7.17. Constrained Control of Uncertain, Time-varying Linear Discrete-Time Systems Subject to Bounded Disturbances

Participants: Sorin Olaru [correspondent], Nam Nguyen [Technion, Israel], Per-Olof Gutman [Technion, Israel], Morten Hovd [NTNU, Trondheim, Norway].

In the paper [26], robust invariance for ellipsoidal sets with respect to uncertain and/or time-varying linear discrete-time systems with bounded additive disturbances is revisited. We provide an extension of an existing invariance condition and propose a novel robust interpolation based control design involving several local unconstrained robust optimal controls. It is shown that at each time instant a quadratic programming problem is solved on-line for the implementation. Proofs of recursive feasibility and input-to-state stability are provided to support the theoretical foundation.

7.18. Explicit robust constrained control for linear systems : analysis, implementation and design based on optimization

Participants: Sorin Olaru [correspondent], Ngoc Anh Nguyen [L2S], Pedro Rodriguez Ayerbe [L2S], Martin Gulan [STUBA, Bratislava].

Piecewise affine (PWA) feedback control laws is relevant for the control of constrained systems, hybrid systems; equally for the approximation of nonlinear control. However, they are associated with serious implementation issues. Motivated from the interest in this class of particular controllers, the thesis of Ngoc Anh Nguyen is mostly related to their analysis and design. The first part of this thesis aims to compute the robustness and fragility margins for a given PWA control law and a linear discrete-time system. More precisely, the robustness margin was defined as the set of linear time-varying systems such that the given PWA control law keeps the trajectories inside a given feasible set. On a different perspective, the fragility margin contains all the admissible variations of the control law coefficients such that the positive invariance of the given feasible set is still guaranteed. The second part of this thesis focuses on inverse optimality problem for the class of PWA controllers. Namely, the goal is to construct an optimization problem whose optimal solution is equivalent to the given PWA function. The methodology is based on convex lifting [33]: an auxiliary 1-dimensional variable which enhances the convexity characterization into recovered optimization problem. Accordingly, if the given PWA function is continuous [35], the optimal solution to this reconstructed optimization problem will be shown to be unique. Otherwise, if the continuity of this given PWA function is not fulfilled, this function will be shown to be one optimal solution to the recovered problem [36]. In view of applications in linear model predictive control (MPC), it was shown that any continuous PWA control law can be obtained by a linear MPC problem with the prediction horizon at most equal to 2 prediction steps [65]. Aside from the theoretical meaning, this result can also be of help to facilitate implementation of PWA control laws by avoiding storing state space partition [66]. Another utility of convex liftings [34] will be shown in the last part of this thesis to be a control Lyapunov function. Accordingly, this convex lifting will be deployed in the so-called robust control design based on convex liftings for linear system affected by bounded additive disturbances and polytopic uncertainties [60]. Both implicit and explicit controllers can be obtained. This method can also guarantee the recursive feasibility and robust stability. However, this control Lyapunov function is only defined over the maximal λ -contractive set for a given subunitary λ which is known to be smaller than the maximal controllable

set. Therefore, an extension of the above method to the N-steps controllable set will be presented. This method is based on a cascade of convex liftings where an auxiliary variable will be used to emulate a Lyapunov function. Namely, this variable will be shown to be non-negative, to strictly decrease for N first steps and to stay at 0 afterwards. Accordingly, robust stability is sought.

7.19. Predictive Control for multi-agent (multi-vehicle) systems

Participants: Sorin Olaru [correspondent], Ionela Prodan [LCIS], Minh Tri Nguyen [L2S], Cristina Stoica [L2S], Silviu Niculescu [L2S], Fernando Fontes [FEUP U. Porto], Joao Sousa [FEUP U. Porto], Fernando Lobo Pereira [FEUP U. Porto], Alexandra Grancharova [U. Sofia, Bulgaria].

We continued a mature line of research on the tracking problems for multi-agent systems. In [75] we presented a series of developments on predictive control for path following via a priori generated trajectory for autonomous aerial vehicles. The strategy partitions itself into offline and runtime procedures with the assumed goal of moving the computationally expensive part into the offline phase and of leaving only tracking decisions to the runtime. First, it will be recalled that differential flatness represents a well-suited tool for generating feasible reference trajectory. Next, an optimization-based control problem which minimizes the tracking error for the nonholonomic system is formulated and further enhanced via path following mechanisms. Finally, possible changes of the selection of sampling times along the path and their impact on the predictive control formulation will be discussed in detail.

On a relatively different framework, in [71] we investigate multiple agents evolving in the same environment with the objective of preservation of a predefined formation. This formation aims to reinforce the safety of the global system and further lighten the supervision task. One of the major issues for this objective is the task assignment problem, which can be formulated in terms of an optimization problem by employing set-theoretic methods. In real time the agents will be steered into the defined formation via task (re)allocation and classical feedback mechanisms. The task assignment calculation is often performed in an offline design stage, without considering the possible variation of the number of agents in the global system. These changes (i.e., including/excluding an agent from a formation) can be regarded as a typical fault, due to some serious damages on the components or due to the operator decision. In this context, our work proposes a new algorithm for the dynamical task assignment formulation of multi-agent systems in view of real-time optimization by including fault detection and isolation capabilities. This algorithm allows to detect whether there is a fault in the global multi-agent system, to isolate the faulty agent and to integrate a recovered/healthy agent. The proposed methods will be illustrated by means of a numerical example with connections to multi-vehicle systems.

7.20. Invariant sets for time-delay systems

Participants: Sorin Olaru [correspondent], Mohammed Laraba [L2S], Silviu Niculescu [L2S].

The characterization of invariant sets for dynamical systems affected by time-delays is a long standing research topic in our group and this year new results have been obtained [52], [51] towards construction of invariant sets in the original state space (also called D-invariant sets) by exploiting the forward mappings. As novelties, the present paper contains a sufficient condition for the existence of ellipsoidal D-contractive sets for dDDEs, and a necessary and sufficient condition for the existence of D-invariant sets in relation to time-varying dDDE stability. Another contribution is the clarification of the relationship between convexity (convex hull operation) and D-invariance. In short, this shown that the convex hull of two D-invariant sets is not D-invariant but the convex hull of a non-convex D-invariant set is D-invariant.

7.21. Biochemical system modeled through a PDE

Participants: Frédéric Mazenc, Abdou Dramé [City Univ. of New York], Peter Wolenski [LSU].

We studied a model of chemostat relying on a Partial Differential Equation. More precisely, we studied in [13] the stability of periodic solutions of distributed parameters biochemical system with periodic input $S_{in}(t)$ (which represent the substrate input concentration). We established that if the function $S_{in}(t)$ is periodic then the system has a periodic solution that possess the robust stability property called input to state stability and this when sufficiently small perturbations are acting on $S_{in}(t)$.

7.22. Modelling of cell dynamics in Acute Myeloid Leukemia

Participants: Catherine Bonnet, Jean Clairambault [BANG project-team], François Delhommeau [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Walid Djema, Emilia Fridman [Tel-Aviv University], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Frédéric Mazenc.

Modelling of Acute Myeloid Leukemia started a few years ago. Starting from a PDE model of hematopoiesis given in [77], we have derived several models of healthy or cancer cell dynamics in hematopoiesis which according to some conditions admit one or two equilibrium points. Often taking profit of the positivity of the system we have derived this year several sufficient (or necessary and sufficient) conditions which ensure stability properties ranging from local asymptotic stability to global exponential stability and obtained, when appropriate, an estimation of a subset of the basin of attraction [14], [40].

7.23. Observability analysis of AC electric machines

Participants: Mohamad Koteich [CentraleSupélec, L2S, Renault], Guillaume Sandou [correspondent], Gilles Duc [CentraleSupélec, L2S], Abdelmalek Maloum [Renault].

High-performance control of electric drives requires an accurate knowledge of the rotor position and/or speed. These mechanical variables are traditionally measured using sensors, which increases the cost and reduces both the robustness and the reliability of the system. This emphasizes the importance of electric drives control without shaft sensors, often referred to as sensorless control : it consists of replacing sensors with a state observer algorithm, that estimates the desired mechanical variables from currents and voltages sensing and based on the system's model. Nevertheless, before designing a state observer, the observability of the system should be examined, that is, it should be checked whether the states to be estimated can be reconstructed, unambiguously, from the input/output signals of the system.

This work addresses the modeling and the observability analysis of electric drives in the view of mechanical sensors removal. Firstly, electrical machines models are elaborated, and it is shown that a unified modeling of alternating current machines is feasible, for the purpose of designing unified control and estimation strategies. The observability of the machines' models is next studied in the view of sensorless control. The local instantaneous observability theory is applied, which enables us to formulate physically insightful analytic conditions that can be easily interpreted and tested in real time. The validity of the observability conditions is confirmed by numerical simulations and experimental data, using an extended Kalman observer. This work contributes to novel outlooks on the sensorless alternating current drives and to a deeper understanding of its properties, in order to develop higher performance estimation techniques in the critical operating conditions (mainly at standstill and/or zerostator-frequency). The concepts introduced throughout this work, such as the equivalent flux and the observability vector, with the obtained results, open new horizons in a domain that seems to become mature enough [48], [76], [46], [49], [45], [15], [16], [91], [92], [47].

7.24. Optimization of Line of Sight controller based on high-level optronic criterion

Participants: Sophie Frasnedo [CentraleSupélec, L2S, Sagem], Guillaume Sandou [correspondent], Gilles Duc [CentraleSupélec, L2S], Philippe Feyel [Sagem], Cédric Chapuis [Sagem].

A method to tune the parameters of the controller of an inertially stabilized platform is proposed. This platform carries an electro-optical system. The image quality is obviously influenced by the movements of the platform: the Line of Sight (LoS) of the imager has to remain fixed in an inertial frame. The more the LoS controller manages to counter the movement of the platform, the better the image quality will be. The motion Modulation Transfer Function (motion MTF) measures the amount of blur brought into the image by the motion of the platform. It represents the contrast over spatial frequencies. Up to now, it has mostly been used as a validation tool for controllers already tuned from derived low level and conservative considerations. The proposed methodology aims to tune LoS controllers using directly the motion MTF as a criterion in the design procedure [42].

7.25. Optimization of controller using bayesian optimization

Participants: Sophie Frasnedo [CentraleSupélec, L2S, Sagem], Julien Bect [CentraleSupélec, L2S], Gilles Duc [CentraleSupélec, L2S], Guillaume Sandou [correspondent], Philippe Feyel [Sagem], Cédric Chapuis [Sagem].

A method to globally optimize the parameters of the controller of an inertially stabilized platform is presented. This platform carries an electro-optical system. The quality of the produced image is obviously influenced by the capacity of the controller to compensate for the unwanted motion of the platform. The motion Modulation Transfer Function (motion MTF) measures the amount of blur brought into the image by those parasite movements. The controller is tuned by minimizing a criterion which includes the motion MTF. However, evaluating this criterion is time-consuming. Using an optimization method that needs numerous evaluations of the criterion is not compatible with industrial constraints. Bayesian optimization methods consist in combining prior information about the criterion and previous evaluation results in order to choose efficiently new evaluation points and reach the global minimizer within a reasonable time. In this paper, a Bayesian approach is used to optimize the motion MTF-based criterion. The results are compared with a local optimization of the same MTF-based criterion, initialized with an acceptable initial point. Similar performances are achieved by the proposed methodology, without requiring an initialization point [41].

7.26. Particle Swarm Optimization based Approach for Model Predictive Control Tuning

Participants: Mohamed Lotfi Derouiche [CentraleSupélec, L2S, Ecole Nationale d'Ingénieur de Tunis], Guillaume Sandou [correspondent], Soufiene Bouallegue [Ecole Nationale d'Ingénieur de Tunis], Joseph Haggège [Ecole Nationale d'Ingénieur de Tunis].

In this work, a new Model Predictive Controller (MPC) parameters tuning strategy is proposed using a perturbed Particle Swarm Optimization (pPSO) approach. This original LabVIEW implementation of this metaheuristic algorithm is firstly validated on some test functions in order to show its efficiency and validity. The optimization results are compared with the standard PSO as well as a LabVIEW implemented Genetic Algorithm (GA) approaches. The parameters tuning problem, i.e. the weighting factors on the output error and input increments of the MPC algorithm, is after that formulated and systematically resolved, using the proposed LabVIEW pPSO algorithm. The case of a Magnetic Levitation (MAGLEV) system is investigated to illustrate the robustness and superiority of the proposed pPSO-based tuning MPC approach. All obtained simulation results, as well as the statistical analysis tests, are compared and discussed in order to improve the effectiveness of the proposed pPSO-based MPC tuning methodology.

7.27. Traffic rescheduling for CBTC train system running in a mixed traffic

Participants: Juliette Pochet [CentraleSupélec, L2S, SNCF], Guillaume Sandou [correspondent], Sylvain Baro [SNCF].

Railway companies need to achieve higher capacities on existing infrastructures such as high density suburban mainlines. Communication based train control (CBTC) systems have been widely deployed on dedicated subway lines. However, deployment on shared rail infrastructure, where CBTC and non CBTC trains run, leads to a mixed positioning and controlling system with different precision levels and restrictions. New performance and complexity issues are to arise. In this work, a method for traffic rescheduling, adapted to a CBTC system running in a mixed traffic, is introduced. A genetic algorithm solves the problem to optimize the cost function. It determines the dwell times and running times of CBTC-equipped trains, taking into account the non-equipped trains planning and fixed-block localization. In addition, reordering can be allowed by modifying the problem constraints. The work is supported by a new simulation tool developed by SNCF and adapted to mixed traffic study. The approach is illustrated with a case study based on a part of an East/West line in Paris region network, proving the ability of the method to find good feasible solutions when delays occur in the traffic.

7.28. Combined Feedback Linearization and MPC for Wind Turbine Power Tracking

Participants: Nicolo Gionfra [CentraleSupélec, L2S], Guillaume Sandou [correspondent], Houria Siguerdjane [CentraleSupélec, L2S], Damien Faille [EDF], Philippe Loevenbruck [EDF].

The problem of controlling a variable-speed-variable-pitch wind turbine in non conventional operating points is addressed. We aim to provide a control architecture for a general active power tracking problem for the turbine's entire operating envelope. The presented control enables to cope with system non linearities while handling state and input constraints, and avoiding singular points. Simulations are carried out based on the CART turbine parameters. Comparative results show that the proposed controller outperforms the classic PI regulator.

DOLPHIN Project-Team

7. New Results

7.1. Benchmarking Numerical Optimizers

Participants: D. Brockhoff, B. Derbel, A. Liefooghe, T.-D. Tran, D. Tušar, T. Tušar (DOLPHIN), O. Ait Elhara, A. Atamna, A. Auger, N. Hansen (TAO team, Inria Saclay), P. Preux (Univ. Lille 3), O. Mersmann, T. Wagner (TU Dortmund University, Germany), B. Bischl (LMU Munich, Germany), Y. Akimoto (Shinshu University, Japan)

In terms of benchmarking numerical optimization algorithms, our research effort went into two different directions. On the one hand, we continued our work on benchmarking single-objective optimization algorithms via the Coco platform in which we started to focus on algorithms for expensive optimization (problems for which only a few function evaluations are affordable). In particular, we benchmarked algorithm variants from the MATSuMoTo library [52], [50] and from the bandits-based global optimizer SOO (Simultaneous optimistic optimization) [33], and organized two workshops at CEC 2015 and GECCO 2015 (see also <http://coco.gforge.inria.fr/>). On the other hand, we started to develop an extension of the Coco platform towards multiobjective optimization and tried to establish the state of the art in single-objective benchmarking (target-based runtimes, data profiles, ...) also in the multi-objective case [30]. At the same time, we proposed a new bi-objective test suite, consisting of 300 well-understood, scalable test problems.

7.2. Handling numeric and temporal data in a local search-based classification algorithm

Participants: J. Jacques, L. Jourdan, C. Dhaenens, M. Vandomme

MOCA-I [20] is a highly efficient classification algorithm, primarily designed for knowledge extraction on large-scale, real-life medical data. This algorithm has been first extended to deal with numeric data [58], [46], through the definition of a model for classification rules on numeric attributes. Several neighborhood operators have been proposed, and compared, as components of the overarching local search metaheuristic guiding the discovery and optimization of these rules. A new model has also been proposed to handle temporal data. This model allows for the inclusion of sequences of events in classification rules, in addition to non-temporal attributes, thus building more informative classifiers. This model, along with various optimizations in the local search process, has been favorably compared to the previous MOCA-I algorithm and other standard classification algorithms. It is now used on real hospital data in order to evaluate its performance in a real environment.

7.3. MO-DYNAMOP

Participants: S. Jacquin, L. Jourdan, E-G. Talbi

The proposed method, MO-DYNAMOP generalize to multi-objective optimization, DYNAMOP, a state of the art optimizer, which was successfully applied to several MO problems. The specificity of this method is to combine aMO dynamic programming (MO-DP) with a MO evolutionary algorithm (MOEA). MO-DYNAMOP is applied to the first stage of the MO-UCP problem including minimization of gas emission. Since the second stage of the problem is now multi-objective, each solution of the first stage problem induces an entire Pareto front of the second stage problem. MO-UCP is solved by assigning an approximation of this Pareto front to each solution of the first stage problem. A comparison study with methods previously proposed in literature is performed. Experiments indicate that MO-DYNAMOP performs considerably better.

7.4. Decomposition-based multi-objective optimization

Participants: Dimo Brockhoff, Bilel Derbel, Arnaud Liefooghe, Gauvain Marquet, El-Ghazali Talbi, Saul Zapotecas-Martinez (external collaborators: Hernan Aguirre and Kiyoshi Tanaka, Shinshu Univ., Japan; Juan Palacios Alonso, Univ. Oviedo, Spain)

MOEA/D is an aggregation-based evolutionary algorithm which has been proved extremely efficient and effective for solving multi-objective optimization problems. It is based on the idea of decomposing the original multi-objective problem into several single-objective subproblems by means of well-defined scalarizing functions. Those single-objective subproblems are solved in a cooperative manner by defining a neighborhood relation between them. This makes MOEA/D particularly interesting when attempting to plug and to leverage single-objective optimizers in a multi-objective setting. For continuous optimization, we investigate in [49] the benefits that MOEA/D can achieve when coupled with CMA-ES, which is believed to be a powerful single-objective optimizer. We rely on the ability of CMA-ES to deal with injected solutions in order to update different covariance matrices with respect to each subproblem defined in MOEA/D. We show that by cooperatively evolving neighboring CMA-ES components, we are able to obtain competitive results for different multi-objective benchmark functions. Moreover, in the combinatorial case, we study in [48] the incorporation of geometric differential evolution (gDE), the discrete generalization of DE, into the MOEA/D framework. We conduct preliminary experiments in order to study the effectiveness of gDE when coupled with MOEA/D. Our results indicate that the proposed approach is highly competitive with respect to the original version of MOEA/D, when solving a combinatorial optimization problem having between two and four objective functions. In [36], we consider a bi-objective scheduling combinatorial problem in which task durations and due-dates are uncertain as a case study for MOEA/D. In particular, we investigate existing variants of MOEA/D and we propose a novel and simple alternative replacement component at the aim of maintaining population diversity. Through extensive experiments, we then provide a comprehensive analysis on the relative performance and the behavior of the considered algorithms. Besides being able to outperform existing MOEA/D variants, as well as the standard NSGA-II algorithm, our investigations provide new insights into the search ability of MOEA/D and highlight new research opportunities for improving its design components. At last, in [32], we propose the first large-scale message passing distributed scheme for parallelizing the computational flow of MOEA/D. We show how synchronicity and workload granularity can impact both quality and computing time, in an extremely fine-grained configuration. We deploy our distributed protocol using a large-scale environment of 128 computing cores. Besides being able to show significant speed-ups while maintaining competitive search quality, our experimental results provide insights into the behavior of the proposed scheme in terms of quality/speed-up trade-offs; thus pushing a step towards the achievement of effective and efficient parallel decomposition-based approaches for large-scale multi-objective optimization.

7.5. Fitness landscape analysis for multi-objective optimization

Participants: F. Daolio, A. Liefooghe (external collaborators: Sébastien Verel, Univ. Littoral Côte d'Opale, France; Hernan Aguirre and Kiyoshi Tanaka, Shinshu Univ., Japan)

Computationally hard multi-objective combinatorial optimization problems are common in practice, and numerous evolutionary multi-objective optimization (EMO) algorithms have been proposed to tackle them. Our aim is to understand which (and how) problem features impact the search performance of such approaches. In [38], we adopt a statistical approach, based on simple and multiple linear regression analysis, to enquire the expected running time of global SEMO with restart for identifying a $(1+\varepsilon)$ -approximation of the Pareto set for small-size enumerable instances. Our analysis provides further insights on the EMO search behavior and on the most important features that characterize the difficulty of an instance for this class of problems and algorithms. In [31], we consider two prototypical dominance-based algorithms: a global EMO strategy using an ergodic variation operator (GSEMO) and a neighborhood-based local search heuristic (PLS). Their respective runtime is estimated on a benchmark of combinatorial problems with tunable ruggedness, objective space dimension, and objective correlation (ρ MNK-landscapes). In other words, benchmark parameters define classes of instances with increasing empirical problem hardness; we enumerate and characterize the search space of small instances. Our study departs from simple performance comparison to systematically analyze

the correlations between runtime and problem features, contrasting their association with search performance within and across instance classes, for both chosen algorithms. A mixed-model approach then allows us to further generalize from the experimental design, supporting a sound assessment of the joint impact of instance features on EMO search performance. Next, in [28], we analyse the behavior and compares the performance of MOEA/D, IBEA using the binary additive epsilon and the hypervolume difference indicators, and A ϵ S ϵ H as representative algorithms of decomposition, indicators, and ϵ -dominance based approaches for many-objective optimization. We use small MNK-landscapes to trace the dynamics of the algorithms generating high-resolution approximations of the Pareto optimal set. Also, we use large MNK-landscapes to analyze their scalability to larger search spaces. At last, in [39], we report an experimental analysis on stochastic local search for approximating the Pareto set of bi-objective unconstrained binary quadratic programming problems. First, we investigate two scalarizing strategies that iteratively identify a high-quality solution for a sequence of sub-problems. Each sub-problem is based on a static or adaptive definition of weighted-sum aggregation coefficients, and is addressed by means of a state-of-the-art single-objective tabu search procedure. Next, we design a Pareto local search that iteratively improves a set of solutions based on a neighborhood structure and on the Pareto dominance relation. At last, we hybridize both classes of algorithms by combining a scalarizing and a Pareto local search in a sequential way. A comprehensive experimental analysis reveals the high performance of the proposed approaches, which substantially improve upon previous best-known solutions. Moreover, the obtained results show the superiority of the hybrid algorithm over non-hybrid ones in terms of solution quality, while requiring a competitive computational cost. In addition, a number of structural properties of the problem instances allow us to explain the main difficulties that the different classes of local search algorithms have to face.

7.6. Fitness Landscape of the Factoradic Representation on the PFSP

Participants: Marie-Eléonore Marmion (external collaborators: Olivier Regnier-Courdert, University of Aberdeen, UK)

Because permutation problems are particularly challenging to model and optimise, the possibility to represent solutions by means of factoradics has recently been investigated, allowing algorithms from other domains to be used. Initial results have shown that methods using factoradics can efficiently explore the search space, but also present difficulties to exploit the best areas. In [57], the fitness landscape of the factoradic representation and one of its simplest operator is studied on the Permutation Flowshop Scheduling Problem (PFSP). The analysis highlights the presence of many local optima and a high ruggedness, which confirms that the factoradic representations is not suited for local search. In addition, comparison with the classic permutation representation establishes that local moves on the factoradic representation are less able to lead to the global optima on the PFSP.

7.7. How Neutrality Helps Multiobjective Local Search Algorithms

Participants: Aymeric Blot, Clarisse Dhaenens, Laetitia Jourdan, Marie-Eléonore Marmion (external collaborators: Hernan Aguirre and Kiyoshi Tanaka, Shinshu Univ., Japan)

We extend the concept of neutrality used in single-objective optimization to the multi-objective context and investigate its effects on the performance of multi-objective dominance-based local search methods [29]. We discuss neutrality in single-objective optimization and fitness assignment in multi-objective algorithms to provide a general definition for neutrality applicable to multi-objective landscapes. We also put forward a definition of neutrality when Pareto dominance is used to compute fitness of solutions. Then, we focus on dedicated local search approaches that have shown good results in multi-objective combinatorial optimization. In such methods, particular attention is paid to the set of solutions selected for exploration, the way the neighborhood is explored, and how the candidate set to update the archive is defined. We investigate the last two of these three important steps from the perspective of neutrality in multi-objective landscapes, propose new strategies that take into account neutrality, and show that exploiting neutrality allows to improve the performance of dominance-based local search methods on bi-objective permutation flowshop scheduling problems. This work is a first step to integrate learning in strategies of local search algorithms.

7.8. Surrogate-assisted multiobjective evolutionary algorithm for fuzzy job shop problems

Participants: E-G. Talbi and Juan José Palacios, Jorge Puente, Camino R. Vela, Inés Gonzalez-Rodríguez (Univ. Oviedo, Spain)

We have considered a job shop scheduling problem with uncertain processing times modelled as triangular fuzzy numbers and propose a multiobjective surrogate assisted evolutionary algorithm to optimise not only the schedule's fuzzy makespan but also the robustness of schedules with respect to different perturbations in the durations. The surrogate model is defined to avoid evaluating the robustness measure for some individuals and estimate it instead based on the robustness values of neighbouring individuals, where neighbour proximity is evaluated based on the similarity of fuzzy makespan values. The experimental results show that by using fitness estimation, it is possible to reach good fitness levels much faster than if all individuals are evaluated.

7.9. Bipartite matching approximation

Participants: F. Dufossé

Bipartite matching is a classical academic problem on bipartite graphs. Many iterative heuristics need an initial approximate matching with linear computational time. We have designed two randomized highly parallelizable algorithms with linear execution time and quality guarantee. The approximation guarantees have been proved to reach respectively 63 demonstrate the speed-ups and validate the applicability and efficiency of these algorithms on general bipartite graphs. Comparisons with the more efficient suboptimal linear algorithms of bipartite matching demonstrate a lower efficiently in average but a similar execution time, and validate the quality guarantee on all experiments. This work has been published in [16].

7.10. Parallel B&B revisited for coprocessors using our new IVM data structure dedicated to permutation problems

Participants: J. Gmys, R. Leroy and N. Melab

This contribution is a joint work with M. Mezmaç and D. Tuytens from University of Mons (UMONS). Solving large permutation Combinatorial Optimization Problems (COPs) using Branch-and-Bound (B&B) algorithms results in the generation of a very large pool of subproblems. Therefore, defining a dedicated data structure is crucial to store and manage efficiently that pool. In the Ph.D thesis of R. Leroy [11], we have proposed an original data structure called Integer-Vector-Matrix (IVM) for permutation COPs based on the factorial number system. Consequently, we have redefined the operators of the B&B algorithm acting on it. For performance evaluation in terms of memory footprint and CPU time usage, we conduct a complexity analysis and an extensive experimentation using the permutation Flow-Shop Scheduling Problem (FSP) as a case study. Compared to the Head-Tail Linked List (LL) data structure often used for parallel B&B as in our work [11], IVM requires up to n times less memory than LL, n being the size of permutations. Moreover, the IVM-based B&B is up to one order of magnitude faster than its LL-based counterpart in managing the pool of subproblems. Another major contribution of this thesis is to revisit parallel B&B for multi-core processors and many-core coprocessors (GPU and MIC) using IVM and LL-based work stealing. Several challenging issues are addressed including work distribution using factoradic-based intervals on multi-core processors, thread/branch divergence and data placement optimization on GPU, and vectorization on Intel Xeon Phi. The contribution and some of its extensions have been published in [40], [18]. An extensive experimental study shows that the IVM-based approach outperforms its LL-based counterpart by a significant margin on multi-core processors as well as on coprocessors.

A major extension of this work has been proposed in [54] and awarded as a best paper consists in offloading all the operators of the B&B algorithm to the GPU. Four interval-based WS strategies have been investigated using IVM. An extensive experimentation allowed us to demonstrate that the GPU-accelerated approach is 5 times faster than its multi-core counterpart.

7.11. large scale heterogeneous parallel B&B based on hybrid work-stealing

Participants: Bilel Derbel, Tuan Trong Vu

In [27], we investigate the design of parallel B&B in large scale heterogeneous compute environments where processing units can be composed of a mixture of multiple shared memory cores, multiple distributed CPUs and multiple GPUs devices. We describe two approaches based on hybrid work-stealing in shared and distributed memory systems, addressing the critical issue of how to map B&B workload with the different levels of parallelism exposed by the target compute platform. We also contribute a throughout large scale experimental study which allows us to derive a comprehensive and fair analysis of the proposed approaches under different system configurations using up to 16 GPUs and up to 512 distributed cores. Our results shed more light on the main challenges one has to face when tackling B&B algorithms while describing efficient techniques to address them. In particular, we are able to obtain linear speed-ups at moderate scales where adaptive load balancing among the heterogeneous compute resources is shown to have a significant impact on performance. At the largest scales, intra-node parallelism and hybrid decentralized load balancing is shown to have a crucial importance in order to alleviate locking issues among shared memory threads and to scale the distributed resources while optimizing communication costs and minimizing idle times.

7.12. A Multi-objective Evolutionary Algorithm for Cloud Platform Reconfiguration

Participants: F. Legillon, N. Melab and E-G. Talbi

This contribution published in [37] is a result of an industrial collaboration with Tasker Cloud services company.

Offers of public IAAS providers are dynamic: new providers enter the market, existing ones change their pricing or improve their offering. The decision on whether and how to improve already deployed platforms, either by reconfiguration or migration to another provider, can be modelled as an NP-hard optimization problem. In this paper, we define a new realistic model for this migration problem, based on a multi-objective Optimization formulation. An evolutionary approach is introduced to tackle the problem, using newly defined specific operators. Experiments are conducted on multiple realistic data-sets, showing that the evolutionary approach is viable to tackle real-size instances in a reasonable amount of time.

7.13. A multi-objective approach for energy-efficient scheduling of large workloads in multicore distributed systems

Participants: E-G. Talbi and B. Dorrnsoro (Univ. Cadiz, Spain), S. Nesmachnow (Universidad de la República, Uruguay), J. Taheri, A. Zomaya (Univ. Sydney, Australia), P. Bouvry (Univ. Luxembourg)

This work proposes a two-level strategy for scheduling large workloads of parallel applications in multicore distributed systems, taking into account the minimization of both the total computation time and the energy consumption of solutions. Nowadays, energy efficiency is of major concern when using large computing systems such as cluster, grid, and cloud computing facilities. In the approach proposed, a combination of higher-level (i.e., between distributed systems) and lower-level (i.e., within each data-center) schedulers are studied for finding efficient mappings of workflows into the resources in order to maximize the quality of service, while reducing the energy required to compute them. The experimental evaluation demonstrates that accurate schedules are computed by using combined list scheduling heuristics (accounting for both problem objectives) in the higher level, and ad-hoc scheduling techniques to take advantage of multicore infrastructures in the lower level. Solutions are also evaluated with two user- and administrator-oriented metrics. Significant improvements are reported on the two problem objectives when compared with traditional load-balancing and round-robin techniques [15].

ECUADOR Project-Team

6. New Results

6.1. AD-adjoints and C dynamic memory management

Participants: Laurent Hascoët, Raphaël Couronné, Sri Hari Krishna Narayanan [Argonne National Lab. (Illinois, USA)], Mathieu Morlighem [University of California at Irvine (USA)].

One of the current frontiers of AD research is the definition of an adjoint AD model that can cope with dynamic memory management. This research is central in our ongoing effort towards adjoint AD of C, and more remotely towards AD of C++. This research is conducted in collaboration with the MCS department of Argonne National Lab. Our partnership is formalized by joint participation in the Inria joint lab JLESC, and partly funded by the Partner University Fund (PUF) of the French embassy in the USA.

Adjoint AD must reproduce in reversed order the control decisions of the original code. In languages such as C, allocation of dynamic memory and pointer management form a significant part of these control decisions. Reproducing memory (de)allocation in reverse means reallocating memory, possibly receiving a different memory chunk. Reproducing pointer addresses in reverse thus require to convert addresses in the former memory chunks into equivalent addresses in the new reallocated chunks. Together with Krishna Narayanan from Argonne, we experiment on real applications to find the most efficient solution to this address conversion problem. We jointly develop a library (called ADMM, ADjoint Memory Management) whose primitives can be used in AD adjoint code to handle this address conversion. Using this library together with Tapenade, we could obtain a correct AD adjoint code of a medium-size industrial code ("Multibody", structural mechanics) that exhibits a typical usage of C pointer arithmetic. This year, the same effort was conducted with the OpenAD AD tool, leading us to an ADMM library less dependent on the particular target AD tool. A joint publication with our colleagues from Argonne is in preparation.

In parallel, we investigate alternative implementation strategies for ADMM, one of which could be to build our own memory (de)allocation mechanism, This should ultimately rely on the standard C library. As a result, management of adjoint memory addresses could be done deeper in the system and therefore with a smaller overhead, at the cost of some additional portability issues.

We pursue our objective of improving reliability of the AD adjoint model for C codes to a similar level as achieved for Fortran. To this end we apply Tapenade to increasingly larger and complex C codes. In addition to the already mentioned "Multibody" application, we initiated differentiation of two new complex applications:

- "BLN" is a code developed by the Inria team ABS, that computes the potential energy of possible conformations of a macromolecule. Its gradient is used to explore the local minima in the energy landscape of these conformations. The AD adjoint of a Fortran implementation of BLN has been built by Tapenade and successfully validated. The adjoint of the C implementation is a challenge that helps us clarify the adjoint AD model that we use in Tapenade. The C version of BLN that we are considering is actually a (partly mechanical) translation of the actual C++ source. This makes this code an even more appealing and challenging test case. This work was mostly conducted by Raphaël Couronné as a part of his summer internship with us.
- "SEISM" is a code developed by Mathieu Morlighem from UC Irvine, jointly with Eric Larour from JPL. This is a glaciology code closely related to the larger "ISSM" code, in C++. One objective, addressed mostly by Mathieu Morlighem, is to clarify recommendations on the C programming style (again very much inspired here from the C++ style) that allows AD to perform better. The other objective, addressed mostly by our team, is to experiment with quite intricate data structures, where Tapenade's static pointer destination analysis is used intensively.

6.2. AD-adjoints of MPI-parallel codes

Participants: Laurent Hascoët, Ala Taftaf, Sri Hari Krishna Narayanan [Argonne National Lab. (Illinois, USA)].

We have a long-standing collaboration with Argonne National Lab on the question of adjoint AD of message-passing parallel codes. We continued joint development of the Adjoinable-MPI library (AMPI) that provides efficient tangent and adjoint AD for MPI-parallel codes, independently of the AD tool used (now AdolC, dco, OpenAD, Tapenade).

During her PhD work, Ala Taftaf is considering the question of checkpointing applied to the AD-adjoint of an MPI-parallel code. Checkpointing is a memory/runtime tradeoff which is essential for adjoint AD of large codes, in particular parallel codes. However, for MPI codes this question has always been addressed by ad-hoc hand manipulations of the differentiated code, and with no formal assurance of correctness. Ala Taftaf is investigating the assumptions implicitly made during past experiments, to clarify and generalize them. On one hand we propose an extension of the adjoint of MPI point-to-point communication primitives, so that the semantics of an adjoint program is preserved for any placement of checkpoints. On the other hand, we propose an alternative extension of these adjoint communications, more efficient but that requires a number of restrictions on the placement of checkpoints. We shall try to provide proof of correctness of these strategies, and in particular demonstrate that they cannot introduce deadlocks. Tradeoffs between the two extensions should be investigated. Ala Taftaf presented her research on “Adjoint-Checkpointing on MPI-parallel codes” at the EuroAD workshop in Paderborn, Germany, december 1-2. A conference article has been submitted to Eccomas 2016 in Crete.

6.3. AD-adjoints of Iterative Processes

Participants: Laurent Hascoët, Ala Taftaf, Sri Hari Krishna Narayanan [Argonne National Lab. (Illinois, USA)], Daniel Goldberg [University of Edinburgh, UK].

Adjoint codes naturally propagate partial gradients backwards from the result of the simulation. However, this uses the data flow of the simulation in reverse order, at a cost that increases with the length of the simulation. In the special case of iterative Fixed-Point loops, only the final converged result should be used: the “initial guess” and the intermediate non-converged states should not be considered by the adjoint calculation, and this remark brings enormous gain in memory use. We selected the strategy proposed by Bruce Christianson [22] and this year we continued its application to medium-size testcases provided by Queen Mary University for the AboutFlow project. We also simplified the user interface provided to trigger this special strategy extension in Tapenade. Ala Taftaf presented her results at the ECCOMAS Eurogen conference in Glasgow [15], september 14-16.

In parallel, we collaborated with Krishna Narayanan from ANL and Dan Goldberg from University of Edinburgh (UK) to implement the same strategy into the OpenAD tool, in view of applying it to a glaciology configuration of the MIT GCM code. A joint article describing the results has been submitted for publication.

6.4. AD-adjoints of large real codes

Participants: Laurent Hascoët, Valérie Pascual, Raphaël Couronné, Fabrice Zaoui [EDF R&D, LNHE].

In collaboration with EDF, Valérie Pascual is applying Tapenade to the hydrographic code “Mascaret”. Both tangent and adjoint differentiated codes have been built and validated. Application of the tangent differentiated Mascaret for data assimilation on two real cases is described in a joint publication [14].

During his summer internship, Raphaël Couronné has applied Tapenade to the MIT “GCM”, a reference code in the Earth Sciences community. We have obtained a valid adjoint for a recommended configuration of this very large Fortran code. This test showed some maturity of the Tapenade tool for Fortran, as it turned out that no modification nor debug of the tool was needed. We are now discussing with the MIT team to schedule further collaboration.

In cooperation with the partners of the FP7 project UMRIDA, the team has assisted Alenia-Aermacchi (Filomena Cariglino and Nicola Ceresola) in the efficient differentiation of their Euler/Navier Stokes code "UNS3D" in tangent mode, dealing in particular with its use of MPI.

The team has assisted Marcin Wyrozowski from Warsaw University of Technology, to apply Tapenade to a CFD software from WUT.

6.5. Resolution of linearised systems and efficiency

Participants: Olivier Allain [Lemma], Gautier Brèthes, Alain Dervieux, Bruno Koobus [Université Montpellier 2], Emmanuelle Itam [Université Montpellier 2], Vincent Lemasle [Lemma], Stephen Wornom [Lemma].

For Fluid Mechanics as well as for Structural Mechanics, an implicit time-advancing is mandatory. It can be applied efficiently if the large systems involved are solved with a good parallel algorithm. In the 90's, a generation of solution algorithms was devised on the basis of Domain Decomposition Methods (DDM). For complex models (compressible flows...), Schwarz DDM were combined with quasi-Newton algorithms such as GMRES. These are for example Restrictive Additive Schwarz (RAS), which is used in our platform AIRONUM. RAS was developed by Cai, Farhat and others. RAS is an ancestor of the widely used class of Newton-Krylov-Schwarz (NKS) algorithms. For hundreds of processors many versions of NKS, and in particular RAS, are almost scalable (convergence rate independent of the number of processors). But scalability vanishes for a medium-large number of processors (thousands). In the ANR ECINADS, coordinated by Ecuador, a Coarse-Grid Deflated RAS was developed: iteration-wise scalability holds for all parts, except for the coarse grid direct solver, which concerns a much smaller problem. Effective Convergence Scalability (ECS) was confirmed up to 2048 processors. Beyond this level the asymptotic complexity of the coarse-grid direct solver becomes predominant and ECS is lost. In other words, with a Coarse-Grid Deflated RAS, the size of the coarse grid problem which is solved by a direct algebraic solver must be limited in order to enjoy ECS. For finer meshes, the coarse system cannot be finer, and the efficiency is lower. It is then natural to consider intermediate meshes on which iterative solvers will be applied. In the ANR MAIDESC, Gautier Brèthes has defined a multi-mesh Full MultiGrid (FMG) algorithm adapted to anisotropic mesh adaptation. In 2015, the method has been extended to MPI-based massive parallelism, in cooperation with the Lemma team for the computation of incompressible flows. As a perspective, our parallel MG can be complemented with the previous version of the solver (deflated RAS) for a higher degree of scalability.

A second issue which we addressed is the use of explicit time advancing. Many unsteady flows have to be computed with explicit time advancing. A single explicit time step is of a low cost and can be highly accurate. Explicit time advancing is mandatory for wave propagation: blast shocks of vortices in wakes. However the meshes used may involve small regions in which the explicit time step should be very small and large regions in which such a small time step is a waste. The family of time-advancing methods in which unsteady phenomena are computed using different time steps in different regions is called the multirate methods. In our cooperation with University of Montpellier, a novel multirate method using cell-agglomeration has been designed and developed in our AIRONUM platform. An article is in preparation. This work takes place in the ANR MAIDESC programme.

6.6. Control of approximation errors

Participants: Gautier Brèthes, Eléonore Gauci, Alain Dervieux, Adrien Loseille [GAMMA team, Inria-Rocquencourt], Frédéric Alauzet [GAMMA team, Inria-Rocquencourt], Stephen Wornom [Lemma], Anca Belme [university of Paris 6].

The study of combination of full multigrid (FMG) with anisotropic mesh adaptation (AA), started with the thesis of Gautier Brèthes, has been published [13].

Further studies of mesh adaptation for viscous flows are currently performed and a journal paper, joint with Inria team Gamma3 and University of Paris 6 (Anca Belme) is in preparation.

An important novelty in mesh adaptation is the norm-oriented AA method. The method relies on the definition of ad hoc correctors. It has been developed in the academic platform “FMG” for elliptic problems. Gautier Brèthes gave several presentations in conferences, a journal article has been submitted. The introduction of the norm-oriented idea considerably amplifies the impact of adjoint-based AA. The applied mathematician and the engineer now have methods when faced to mesh adaptation for the simulation of a complex PDE system, since they can specify which error norm level they wish, and for which norm [12], [16]. Another version is developed jointly with Inria team Gamma3 for the compressible Euler model [19].

A cooperation has started between Gautier Brèthes et Thierry Coupez (Ecole Centrale de Nantes) concerning discrete metrics. This takes place in the ANR MAIDESC program. An article is in preparation.

Éléonore Gauci started last year a thesis (co-advised by Frédéric Alauzet) on the study of norm-oriented criteria for CFD and coupled CSM-CFD systems. She gave a presentation at the “Coupled Problems” symposium.

Post-doc Guilherme Cunha did a study (in cooperation with Lemma) on the combination of mesh adaptation and coefficient identification for unsteady phenomena.

The theoretical studies are supported by an ANR project MAIDESC coordinated by ECUADOR and Gamma3, which deals with meshes for interfaces, third-order accuracy, meshes for boundary layers, and curved meshes.

CFD application are supported by the European FP7 project UMRIDA which deals with the application of AA to approximation error modelling and control.

6.7. Turbulence models

Participants: Alain Dervieux, Bruno Koobus [University of Montpellier 2], Emmanuelle Itam [University of Montpellier 2], Marianna Braza [CNRS-IMFT at Toulouse], Stephen Wornom [Lemma], Bruno Sainte-Rose [Lemma].

The purpose of our work in hybrid RANS/LES is to develop new approaches for industrial applications of LES-based analyses. In the applications targetted (aeronautics, hydraulics), the Reynolds number can be as high as several tenth millions, far too high for pure LES models. However, certain regions in the flow can be better predicted with LES than with usual statistical RANS (Reynolds averaged Navier-Stokes) models. These are mainly vortical separated regions as assumed in one of the most popular hybrid model, the hybrid Detached Eddy Simulation model. Here, “hybrid” means that a blending is applied between LES and RANS. An important difference between a real life flow and a wind tunnel or basin is that the turbulence of the flow upstream of each body is not well known.

This year, we have continued the evaluation of a dynamic formulation of Piomelli-Germano type for the Variational-multiscale model. We have also modified the integration of the boundary layer by adding the so-called Menter correction imposing the Bradshaw law. We have studied these improvements on multiple-body flows. An emblematic case is the interaction between two parallel cylinders, one being in the wake of the other. A flow around a space probe at high Reynolds number is also studied [18], [17].

GAMMA3 Project-Team

5. New Results

5.1. Serendipity and reduced elements

Participants: Paul Louis George [correspondant], Houman Borouchaki, Nicolas Barral.

We give a method to constructing Serendipity elements for quads and hexes with full symmetry properties and indicate the reading of their shape functions. We show that, since the degree 5, the Serendipity elements are no longer symmetric but we propose a method resulting in a Lagrange element of degree 5 with full symmetry properties after adding an adequate number of additional nodes.

On the other hand, we show how to guarantee the geometric validity of a given curved element (seen as a patch) of a mesh. This is achieved after writing the patch in a Bézier setting (Bernstein polynomials and control points). In addition, we discuss the case of patch derived from a transfinite interpolation and it is proved that only some of them are Serendipity elements indeed, we return to the same elements as above

We also give a method to constructing Lagrange Serendipity (or reduced) simplices with a detailed description of the triangles of degree 3 and 4. We indicate that higher order triangles are not candidate apart if we impose a restricted polynomial space. We show that a tetrahedron of degree 3 is a candidate while high order elements are not candidate even if a restriction in the polynomial space is considered. In addition, we propose a method for the validation of such elements, in a given mesh, where the validation means the positiveness of the jacobian.

5.2. Validity of transfinite and Bézier-Serendipity patches

Participants: Paul Louis George [correspondant], Houman Borouchaki, Nicolas Barral.

We define generalized transfinite patches for quads and hexes with full symmetry properties. We give a way of constructing those patches by considering the Bézier setting using linear combinations of tensor-product patches of various degree. Those patches are exactly the Bézier-Serendipity patches recently introduced

ASfor reduced quadrilateral patches, we introduce the so called "Bézier-Serendip" patches. After some recalls about standard Bézier patches, we propose a method to constructing those reduced patches. The corresponding Bernstein polynomials are written by means of linear combinations of the standard Bernstein polynomials. We give a full description of the patches of degree 2, 3, 4 and 5. Since degree 5, the location of the control points is no longer symmetric and to remedy this problem, we propose adding a number of control points which results in *extended* Bézier-Serendip patches. Those reduced patches are in the Bézier framework what the Serendipity elements are in the finite element framework.

A technical report and a paper have been published [16].

5.3. Meshing Strategies and the Impact of Finite Element Quality on the Velocity Field in Fractured Media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

For calculating flow in a fracture network, the mixed hybrid finite element (MHFE) method is a method of choice as it yields a symmetric, positive definite linear system. However, a drawback to this method is its sensitivity to bad aspect ratio elements. For poor-quality triangles, elementary matrices are ill-conditioned, and inconsistent velocity vectors are obtained by inverting these local matrices. In this work, different strategies have been proposed for better reconstruction of the velocity field.

5.4. Automatic Mesh Generation of Multiface Models on Multicore Processors

Participant: Patrick Laug [correspondant].

This work started in September 2014, as part of a sabbatical year at Polytechnique Montréal. In a previous study, a parallel version of an indirect approach for meshing composite surfaces – also called multiface models – was developed. However, this methodology could be inefficient in practice, as the memory management of most existing CAD (computer aided design) systems use static global caches to save information. In a first approach, CAD queries are fully parallelized, using the Pirate library from Polytechnique Montréal (this library provides a set of C++ classes that implement STEP-compliant B-Rep geometric and topological entities, as well as classes to represent meshes and solutions). In a second approach, the CAD system is completely disconnected from the mesh generator, using a discrete geometric support.

5.5. Applications du maillage et développements de méthodes avancées pour la cryptographie

Participants: Thomas Grosge [correspondant], Dominique Barchiesi, Michael François.

L'utilisation des nombres (pseudo)-aléatoires a pris une dimension importante ces dernières décennies. De nombreuses applications dans le domaine des télécommunications, de la cryptographie, des simulations numériques ou encore des jeux de hasard, ont contribué au développement et à l'usage de ces nombres. Les méthodes utilisées pour la génération de tels nombres (pseudo)-aléatoires proviennent de deux types de processus : physique et algorithmique. Ce projet de recherche a donc pour objectif principal le développement de nouveaux procédés de génération de clés de chiffrement, dits "exotiques", basés sur des processus physiques, multi-échelles, multi-domaines assurant un niveau élevé de sécurité. Deux classes de générateurs basés sur des principes de mesures physiques et des processus mathématiques ont été développés.

La première classe de générateurs exploite la réponse d'un système physique servant de source pour la génération des séquences aléatoires. Cette classe utilise aussi bien des résultats de simulation que des résultats de mesures interférométriques pour produire des séquences de nombres aléatoires. L'application du maillage adaptatif sert au contrôle de l'erreur sur la solution des champs physiques (simulés ou mesurés). A partir de ces cartes physiques, un maillage avec estimateur d'erreur sur l'entropie du système est appliqué. Celui-ci permet de redistribuer les positions spatiales des noeuds. L'étude (locale) de la réduction d'entropie des clés tout au long de la chaîne de création et l'étude (globale) de l'entropie de l'espace des clés générées sont réalisées à partir de tests statistiques.

La seconde classe de générateurs porte sur le développement de méthodes avancées et est basée sur l'exploitation de fonctions chaotiques en utilisant les sorties de ces fonctions comme indice de permutation sur un vecteur initial. Ce projet s'intéresse également aux systèmes de chiffrement pour la protection des données et deux algorithmes de chiffrement d'images utilisant des fonctions chaotiques sont développés et analysés. Ces Algorithmes utilisent un processus de permutation-substitution sur les bits de l'image originale. Une analyse statistique approfondie confirme la pertinence des cryptosystèmes développés.

5.6. Développement de méthodes avancées et maillages appliqués à l'étude de la nanomorphologie des nanotubes-fils en suspension liquide

Participants: Thomas Grosge [correspondant], Dominique Barchiesi, Abel Cherouat, Houman Borouchaki, Laurence Giraud-Moreau, Anis Chaari.

Ce projet de recherche (NANOMORPH) a pour objet principal le développement et la mise au point d'une instrumentation optique pour déterminer la distribution en tailles et le coefficient de forme de nanofils (NF) ou de nanotubes (NT) en suspension dans un écoulement. Au cours de ce projet, deux types de techniques optiques complémentaires sont développées. La première, basée sur la diffusion statique de la lumière, nécessite d'étudier au préalable la physico-chimie de la dispersion, la stabilisation et l'orientation des nanofils dans les milieux d'étude. La seconde méthode, basée sur une méthode opto-photothermique pulsée, nécessite en sus,

la modélisation de l'interaction laser/nanofils, ainsi que l'étude des phénomènes multiphysiques induits par ce processus. L'implication de l'équipe-projet GAMMA3 concerne principalement la simulation multiphysique de l'interaction laser-nanofils et l'évolution temporelle des bulles et leurs formations. L'une des principales difficultés de ces problématiques est que la géométrie du domaine est variable (à la fois au sens géométrique et topologique). Ces simulations ne peuvent donc être réalisées que dans un schéma adaptatif de calcul nécessitant le remaillage tridimensionnel mobile, déformable avec topologie variable du domaine (formation et évolution des bulles au cours du temps et de l'espace).

5.7. Applications du maillage à des problèmes multi-physiques, développement de méthodes de résolutions avancées et modélisation électromagnétique-thermique-mécanique à l'échelle mesoscopique

Participants: Dominique Barchiesi [correspondant], Abel Cherouat, Thomas Grosgees, Houman Borouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari, Fadhil Mezghani.

Le contrôle et l'adaptation du maillage lors de la résolution de problèmes couplés ou/et non linéaires reste un problème ouvert et fortement dépendant du type de couplage physique entre les EDP à résoudre. Notre objectif est de développer des modèles stables afin de calculer les dilatations induites par l'absorption d'énergie électromagnétique, par des structures matérielles inférieures au micron. Les structures étudiées sont en particulier des nanoparticules métalliques en condition de résonance plasmon. Dans ce cas, un maximum d'énergie absorbée est attendu, accompagné d'un maximum d'élévation de température et de dilatation. Il faut en particulier développer des modèles permettant de simuler le comportement multiphysique de particules de formes quelconques, pour une gamme de fréquences du laser d'éclairage assez étendue afin d'obtenir une étude spectroscopique de la température et de la dilatation. L'objectif intermédiaire est de pouvoir quantifier la dilatation en fonction de la puissance laser incidente. Le calcul doit donc être dimensionné et permettre finalement des applications dans les domaines des capteurs et de l'ingénierie biomédicale. En effet, ces nanoparticules métalliques sont utilisées à la fois pour le traitement des cancers superficiels par nécrose de tumeur sous éclairage adéquat, dans la fenêtres de transparence cellulaire. Déposées sur un substrat de verre, ces nanoparticules permettent de construire des capteurs utilisant la résonance plasmon pour être plus sensibles (voir projet européen *Nanoantenna* et l'activité génération de nombres aléatoires). Cependant, dans les deux cas, il est nécessaire, en environnement complexe de déterminer la température locale, voire la dilatation de ces nanoparticules, pouvant conduire à un désaccord du capteur, la résonance plasmon étant très sensible aux paramètres géométriques et matériels des nanostructures. Dans ce sens, l'étude permet d'aller plus loin que la "simple" interaction électromagnétique avec la matière du projet européen *Nanoantenna*.

Le travail de l'année 2014 a constitué en la poursuite de l'étude des spécificités de ce type de problème multiphysique pour des structures de forme simple et la mise en place de fonctions test, de référence, pour les développements de maillage adaptatifs pour les modèles multiphysiques éléments finis. Nous espérons pouvoir proposer un projet ANR couplant les points de vue microscopiques et macroscopiques dans les deux années qui viennent.

5.8. Visualization and modification of high-order curved meshes

Participants: Alexis Loyer, Adrien Loseille [correspondant].

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh and hybrid meshes has been addressed. The software VIZIR is freely available at <https://www.rocq.inria.fr/gamma/gamma/vizir/>.

5.9. Mesh adaptive ALE numerical simulation

Participants: Frédéric Alauzet [correspondant], Nicolas Barral, Adrien Loseille.

Running highly accurate numerical simulations with moving geometries is still a challenge today due to their prohibitive cost in CPU time. Using anisotropic mesh adaptation is one way to drastically reduce the size of the problem and to reach the desired accuracy. Previously, we have developed an ALE formulation using mesh connectivity change in order to achieve any complex displacement. Then, this method has been coupled with the unsteady anisotropic mesh adaptation using the fixed-point algorithm. The key point of this work is the use of an ALE metric that takes into account the mesh motion in the metric field definition.

5.10. Mesh adaptation for Navier-Stokes Equations

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

Adaptive simulations for Navier-Stokes equations require to propose accurate error estimates and design robust mesh adaptation algorithms (for boundary layers).

For error estimates, we design new estimates suited to accurately capture the speed profile in the boundary layers. For mesh adaptation, we design a new method to generate structured boundary layer meshes which are mandatory to accurately compute compressible flows a high Reynolds number (several millions). It couple the specification of the optimal boundary layer from the geometry boundary and moving mesh techniques to extrude the boundary layer in an already existing mesh. The main advantage of this approach is its robustness, *i.e.*, at each step of the algorithm we have always a valid mesh [23].

5.11. Adaptive multigrid strategies

Participants: Frédéric Alauzet [correspondant], Victorien Menier, Adrien Loseille.

Multigrid is a well known technique used to accelerate the convergence of linear system solutions. Using a multigrid strategy to solve non-linear problems improves the robustness and the convergence of each Newton step, the accelerating overall the whole process. In particular, larger time step can be considered. This of main importance when solving turbulent Navier-Stokes equations on complex geometries. First, we developed the classical multigrid method on non-nested meshes. Then, we have pointed out the similarity between the Full MultiGrid (FMG) algorithm and the mesh adaptation algorithm. We have proposed a new Adaptive Full MultiGrid algorithm which improve the overall robustness of the adaptive process and its overall efficiency [23].

5.12. Metric-orthogonal and metric-aligned mesh adaptation

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

A new algorithm to derive adaptive meshes has been introduced through new cavity-based algorithms. It allows to generate anisotropic surface and volume mesh that are aligned along the eigenvector directions. This allows us to improv the quality of the meshes and to deal naturally with boundary layer mesh generation.

5.13. Parallel mesh adaptation

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

We devise a strategy in order to generate large-size adapted anisotropic meshes $O(10^8 - 10^9)$ as required in many fields of application in scientific computing. We target moderate scale parallel computational resources as typically found in R&D units where the number of cores ranges in $O(10^2 - 10^3)$. Both distributed and shared memory architectures are handled. Our strategy is based on typical domain splitting algorithm to remesh the partitions in parallel. Both the volume and the surface mesh are adapted simultaneously and the efficiency of the method is independent of the complexity of the geometry. The originality of the method relies on (i) a metric-based static load-balancing, (ii) dedicated mesh partitioning techniques to (re)split the (complex) interfaces meshes, (iii) anisotropic Delaunay cavity to define the interface meshes, (iv) a fast, robust and generic sequential cavity-based mesh modification kernel, and (v) out-of-core storing of completing parts to reduce the memory footprint. We are able to generate (uniform, isotropic and anisotropic) meshes with more than 1 billion tetrahedra in less than 20 minutes on 120 cores.

5.14. Unsteady adjoint computation on dynamic meshes

Participants: Eléonore Gauci, Frédéric Alauzet [correspondant].

Adjoint formulations for unsteady problems are less common in unsteady methodologies due to the extra complexity inherent in the numerical solution and storage but these methods are a great option in engineering because it takes more into account the cost function we want to minimize. Moreover the engineering applications involve moving elements and this motion must be taken into account by the governing flow equations. We develop a model of unsteady adjoint solver on moving mesh problems. The derivation of the adjoint formulation based on the ALE form of the equations requires consideration of the dynamic meshes. Our model takes into account the DGCL.

5.15. Line solver for efficient stiff parse system resolution

Participants: Loïc Frazza, Frédéric Alauzet [correspondant].

Afin d'accélérer la résolution des problèmes raides, un line-solver a été développé. Cette méthode extrait tout d'abord des lignes dans le maillage du problème selon des critères géométriques ou physiques. Le problème peut alors être résolu exactement le long de ces lignes à moindre coût. Cette méthode est particulièrement bien adaptée aux cas où l'information se propage selon une direction privilégiée tels que les chocs, les couches limites ou les sillages. Ces cas sont généralement associés à des maillages très étirés ce qui conduit à des problèmes raides mais quasi-unidimensionnels. Ils peuvent donc être résolus efficacement par un line-solver, réduisant ainsi les temps de calculs tout en gagnant en robustesse.

5.16. Error estimate for high-order solution field

Participants: Olivier Coulaud, Adrien Loseille [correspondant].

Afin de produire des solveurs d'ordre élevé, et ainsi répondre aux exigences inhérentes à la résolution de problèmes physiques complexes, nous développons une méthode d'adaptation de maillage d'ordre élevé. Celle-ci est basée sur le contrôle par une métrique de l'erreur d'interpolation induite par le maillage du domaine. Plus précisément, pour une solution donnée, l'erreur d'interpolation d'ordre k est paramétrée par la différentielle k^{e} de cette solution, et le problème se réduit à trouver la plus grande ellipse incluse dans une ligne de niveau de cette différentielle. S'il reste encore quelques difficultés techniques à résoudre avant l'exploitation numérique de notre méthode, les résultats sont très encourageants, tant en terme d'optimalité de la métrique obtenue que de temps de calcul. Il n'y a que peu de doutes sur le fait que ce projet aboutisse prochainement.

GECO Project-Team

6. New Results

6.1. New results: geometric control

Let us list some new results in sub-Riemannian geometry and hypoelliptic diffusion obtained by GECO's members.

- In [12] and [20] we study the sub-Finsler geometry as a time-optimal control problem. In particular, we consider non-smooth and non-strictly convex sub-Finsler structures associated with the Heisenberg, Grushin, and Martinet distributions. Motivated by problems in geometric group theory, we characterize extremal curves, discuss their optimality, and calculate the metric spheres, proving their Euclidean rectifiability.
- In [18] we compare different notions of curvature on contact sub-Riemannian manifolds. In particular we introduce canonical curvatures as the coefficients of the sub-Riemannian Jacobi equation. The main result is that all these coefficients are encoded in the asymptotic expansion of the horizontal derivatives of the sub-Riemannian distance. We explicitly compute their expressions in terms of the standard tensors of contact geometry. As an application of these results, we obtain a sub-Riemannian version of the Bonnet-Myers theorem that applies to any contact manifold.
- In sub-Riemannian geometry the coefficients of the Jacobi equation define curvature-like invariants. We show in [21] that these coefficients can be interpreted as the curvature of a canonical Ehresmann connection associated to the metric, first introduced by Zelenko and Li. We show why this connection is naturally nonlinear, and we discuss some of its properties.
- On a sub-Riemannian manifold we define in [22] two type of Laplacians. The macroscopic Laplacian, as the divergence of the horizontal gradient, once a volume is fixed, and the microscopic Laplacian, as the operator associated with a geodesic random walk. We consider a general class of random walks, where all sub-Riemannian geodesics are taken in account. This operator depends only on the choice of a complement to the sub-Riemannian distribution. We address the problem of equivalence of the two operators. This problem is interesting since, on equiregular sub-Riemannian manifolds, there is always an intrinsic volume (e.g. Popp's one) but not a canonical choice of complement. The result depends heavily on the type of structure under investigation: we describe the relationship between the two approaches in the case of contact structures, Carnot groups, quasi-contact structures.
- In [2] we study 3D almost-Riemannian manifolds, that is, generalized Riemannian manifolds defined locally by 3 vector fields that play the role of an orthonormal frame, but could become collinear on some singular set. Almost-Riemannian manifolds were deeply studied in dimension 2. In this paper we start the study of the 3D case which appear to be reacher with respect to the 2D case, due to the presence of abnormal extremals which define a field of directions on the singular set. We study the type of singularities of the metric that could appear generically, we construct local normal forms and we study abnormal extremals. We then study the nilpotent approximation and the structure of the corresponding small spheres. We finally give some preliminary results about heat diffusion on such manifolds.

New results on motion planning are the following.

- In [7] (written while D. Prandi was PhD student in the team) we study the complexity of the motion planning problem for control-affine systems. Such complexities are already defined and rather well-understood in the particular case of nonholonomic (or sub-Riemannian) systems. Our aim is to generalize these notions and results to systems with a drift. Accordingly, we present various definitions of complexity, as functions of the curve that is approximated, and of the precision of the approximation. Due to the lack of time-rescaling invariance of these systems, we consider geometric

and parametrized curves separately. Then, we give some asymptotic estimates for these quantities. As a byproduct, we are able to treat the long-time local controllability problem, giving quantitative estimates on the cost of stabilizing the system near a non-equilibrium point of the drift.

- In [11] and [1] we propose new conditions guaranteeing that the trajectories of a mechanical control system can track any curve on the configuration manifold. We focus on systems that can be represented as forced affine connection control systems and we generalize the sufficient conditions for tracking known in the literature. The new results are proved by a combination of averaging procedures by highly oscillating controls and the notion of kinematic reduction.
- In
- [17] we introduce the concept of Developmental Partial Differential Equation (DPDE), which consists of a Partial Differential Equation (PDE) on a time-varying manifold with complete coupling between the PDE and the manifold's evolution. In other words, the manifold's evolution depends on the solution to the PDE, and vice versa the differential operator of the PDE depends on the manifold's geometry. DPDE is used to study a diffusion equation with source on a growing surface whose growth depends on the intensity of the diffused quantity. The surface may, for instance, represent the membrane of an egg chamber and the diffused quantity a protein activating a signaling pathway leading to growth. Our main objective is to show controllability of the surface shape using a fixed source with variable intensity for the diffusion. More specifically, we look for a control driving a symmetric manifold shape to any other symmetric shape in a given time interval. For the diffusion we take directly the Laplace-Beltrami operator of the surface, while the surface growth is assumed to be equal to the value of the diffused quantity. We introduce a theoretical framework, provide approximate controllability and show numerical results. Future applications include a specific model for the oogenesis of *Drosophila melanogaster*.

6.2. New results: quantum control

New results have been obtained for the control of the bilinear Schrödinger equation.

- In [4] we study the so-called spin-boson system, namely a two-level system in interaction with a distinguished mode of a quantized bosonic field. We give a brief description of the controlled Rabi and Jaynes–Cummings models and we discuss their appearance in the mathematics and physics literature. We then study the controllability of the Rabi model when the control is an external field acting on the bosonic part. Applying geometric control techniques to the Galerkin approximation and using perturbation theory to guarantee non-resonance of the spectrum of the drift operator, we prove approximate controllability of the system, for almost every value of the interaction parameter.
- The main result in [9] is the approximate controllability of a bilinear Schrödinger equation modeling a system of two ions trapped in a cavity. A new spectral decoupling technique is introduced, which allows to analyze the controllability of the infinite-dimensional system through finite-dimensional considerations. The controllability of a simplified version of the model has been obtained in [16].
- In [13] and [3] we study the controllability of a closed control-affine quantum system driven by two or more external fields. We provide a sufficient condition for controllability in terms of existence of conical intersections between eigenvalues of the Hamiltonian in dependence of the controls seen as parameters. Such spectral condition is structurally stable in the case of three controls or in the case of two controls when the Hamiltonian is real. The spectral condition appears naturally in the adiabatic control framework and yields approximate controllability in the infinite-dimensional case. In the finite-dimensional case it implies that the system is Lie-bracket generating when lifted to the group of unitary transformations, and in particular that it is exactly controllable. Hence, Lie algebraic conditions are deduced from purely spectral properties. We conclude the analysis by proving that approximate and exact controllability are equivalent properties for general finite-dimensional quantum systems.

In [26], written with the members of the European project QUAINT, state-of-the-art quantum control techniques are reviewed and put into perspective by a consortium uniting expertise in optimal control theory and applications to spectroscopy, imaging, quantum dynamics of closed and open systems. Key challenges are addressed and a roadmap to future developments is sketched.

6.3. New results: neurophysiology

In [27] we present a new version of the image inpainting algorithm that GECO developed in the recent years. This new version is called the Averaging and Hypoelliptic Evolution (AHE) algorithm, and is based upon a semi-discrete variation of the Citti–Petitot–Sarti model of the primary visual cortex V1. In particular, we focus on reconstructing highly corrupted images (i.e. where more than the 80% of the image is missing).

6.4. New results: switched systems

- In [5] we consider a continuous-time linear switched system on \mathbb{R}^n associated with a compact convex set of matrices. When it is irreducible and its largest Lyapunov exponent is zero there always exists a Barabanov norm associated with the system. We look at two types of issues: (a) properties of Barabanov norms such as uniqueness up to homogeneity and strict convexity; (b) asymptotic behaviour of the extremal solutions of the linear switched system. Regarding Issue (a), we provide partial answers and propose four related open problems. As for Issue (b), we establish, when $n = 3$, a Poincaré–Bendixson theorem under a regularity assumption on the set of matrices. We then revisit a noteworthy result of N.E. Barabanov describing the asymptotic behaviour of linear switched system on \mathbb{R}^3 associated with a pair of Hurwitz matrices $\{A, A + bc^T\}$.
- Motivated by an open problem posed by J.P. Hespanha, in [23] we extend the notion of Barabanov norm and extremal trajectory to classes of switching signals that are not closed under concatenation. We use these tools to prove that the finiteness of the L_2 -gain is equivalent, for a large set of switched linear control systems, to the condition that the generalized spectral radius associated with any minimal realization of the original switched system is smaller than one.
- In [14] in the totally observed case and in [23] in the general case, we answer an open problem posed by J.P. Hespanha in 2003. We first extend the notion of Barabanov norm and extremal trajectory to classes of switching signals that are not closed under concatenation. We use these tools to prove that the finiteness of the L_2 -gain is equivalent, for a large set of switched linear control systems, to the condition that the generalized spectral radius associated with any minimal realization of the original switched system is smaller than one.
- In [24] we address the stability of non-autonomous difference equations by providing a suitable representation of the solution at time t in terms of the initial condition and time-dependent matrix coefficients. This enables us to characterize the asymptotic behavior of solutions in terms of that of such coefficients. As a consequence, we obtain necessary and sufficient stability criteria for non-autonomous linear difference equations. In the case of difference equations with arbitrary switching, we obtain a generalization of the well-known criterion for autonomous systems due to Hale and Silkowski, which, as the latter, is delay-independent. These results are applied to transport and wave propagation on networks. In particular, we show that the wave equation on a network with arbitrarily switching damping at external vertices is exponentially stable if and only if the network is a tree and the damping is bounded away from zero at all external vertices but one.
- For linear systems in continuous time with random switching, we characterize in [25] the Lyapunov exponents using the Multiplicative Ergodic Theorem for an associated system in discrete time. An application to control systems shows that here a controllability condition implies that arbitrary exponential decay rates for almost sure stabilization can be obtained.

A result related to switched system is the one obtained in [6] and [15]: we study the stability of linear time-varying delay differential equations where the delay enters as a switching parameter. In [6] we give a collection of converse Lyapunov–Krasovskii theorems for uncertain retarded differential equations. We show that the

existence of a weakly-degenerate Lyapunov–Krasovskii functional is a necessary and sufficient condition for the global exponential stability of linear retarded functional differential equations. In [15] the fundamental question that we consider is the following: assuming that every individual (constant-delay) subsystem is exponentially stable, can we characterize the cases when the system is not exponentially stable? This is nothing else than the so-called Markus-Yamabe instability and we give new conditions ensuring it.

GEOSTAT Project-Team

7. New Results

7.1. Super-resolution, multiscale data fusion and complex dynamics in Earth Observation and Universe Sciences

Participants: Hussein Yahia, Nicolas Brodu, Guillaume Attuel, Sylvain Bontemps, Nicola Schneider, Camila Artana, Dharmendra Singh, Joel Sudre, Véronique Garçon, Christine Provost, Anass El Aouni, Oriol Pont, Khalid Daoudi, Ayoub Tamim, Akankhsa Garg, Frédéric Frappart, Luc Bourrel.

In these thematics the following research is started or continued:

- Super-resolution and data fusion in Earth Observation. Important results obtained in validation either in ocean dynamics or partial pCO_2 pressures in ocean/atmosphere exchanges, coastal upwelling.
- Development of a new super-resolution model for multispectral images, demonstration on MODIS (NASA) and Sentinel-2 (ESA) data.
- Adaptive optics.
- Starting of a strong collaboration with Labroatoire d'Astrophysique de Bordeaux on the dynamics of galactic clouds.
- Supervised classification of ground terrain through multispectral imagery (with OPTIC associated team).
- Anomaly detection in SAR images, application to flood monitoring in Equator.
- Starting of a project on dune monitoring.

Publications: [21], [19], [29], [22], [25], [18] [A. Tamim's PhD HAL link](#), [IEEE TGRS article on AO \[21\]](#).

7.2. Characterization of underlying stochastic dynamic of the cardiac muscle under fibrillation: singularity analysis and modeling

Participants: Guillaume Attuel, Binbin Xu, Oriol Pont, Hussein Yahia.

Signals of heart electrical activity obtained through invasive measurements show properties not compatible with the purely excitable nature of cellular dynamics. We have developed a synaptic perturbation model of that dynamics showing good properties. Perturbations propagate an inter-cell desynchronization formally like diffusion-coupled chaotic maps. The model enters the universality class of directed propagative fronts of the type random branching or directed polymer in a disordered medium in 1+1D and pinning-depinning contact lines in 1+2D. In the continuum limit, the universality class is supposed to be the one of KPZ (Kardar Parisi Zhang) or VM (Voter Model). This is a change in paradigm for the description of cardiac dynamics. We make use of this hypothesis to characterize precisely the state of the substratum through appropriate signal analysis, with the goal of being able to distinguish between different states or types in the pathology. We are involved in a technological transfert on this activity since summer 2015.

Publication: [26].

7.3. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

Participant: Binbin Xu.

Background: The animal models (in vitro or in vivo) provide an excellent tool to study heart diseases, among them the arrhythmia remains one of the most active research subjects. It can be induced or treated by drugs, electrical stimulation, hypothermia etc. **Problems:** However, the inducing or treating effects in cardiac culture often happened long after the initial applications or in some relatively short time windows. So, it is necessary to capture and classify rapidly the signal change. Human-assisted monitoring is time-consuming and less efficient. An automatic classification method for real-time use would be useful and necessary. **Methods:** Since electrocardiological signals are features by repetitive or similar patterns reflecting the intrinsic information about the patient (or culture), analyzing these patterns could help not only to monitor the status's change but also to evaluate/explore the physiologic control mechanisms. Methods based on complexity analysis are of considerable interest in this case. **Aims:** Compare different complexity analysis methods in order to find the most appropriate ones to discriminate the normal cardiac signals from arrhythmic ones acquired from a cardiac cell culture in vitro. The selected features are then used by a SVM classifier.

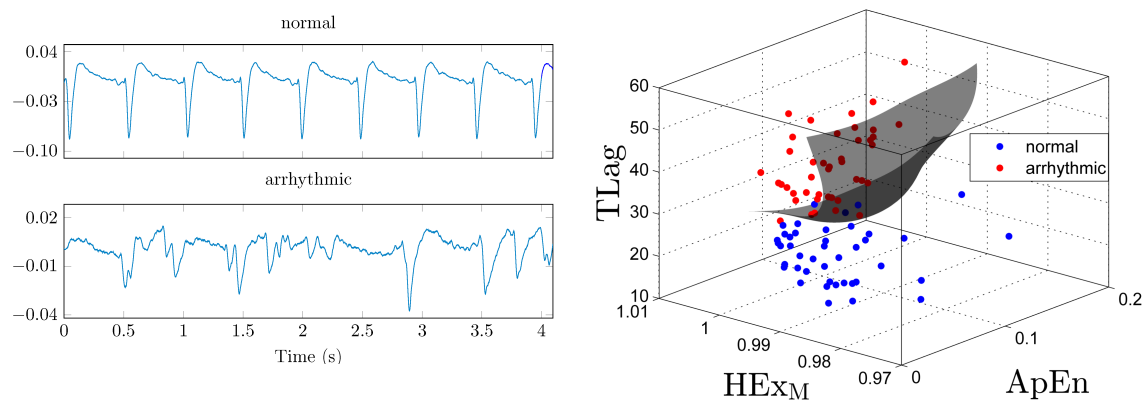


Figure 1. Left: Electrical field potential (EFP) of in vitro cardiac culture on Multielectrodes Array; Right: SVM classification of normal and arrhythmic EFP signals.

Results: Among the six complexity analysis methods, Time Lagging (TLag) method allowed obtaining the best discrimination index (normal vs. arrhythmic, p -value, $9e-23$). The proposed Modified Hurst Exponent (HEX_M) showed better performance than original Hurst Exponent with well-improved p -value (from 0.019 to $2e-9$). The Approximate Entropy (ApEn), Sample Entropy (SampEn) and Detrended Fluctuation Analysis gave good discrimination ratio but with larger p -values (at order 10^{-3}). Combination of TLag, HEX_M and ApEn can provide a more robust classifier and allow monitoring and classifying in an automatic way the electrical activities' changes in the cardiac cultures.

Publication: [28].

7.4. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

Participant: Binbin Xu.

Physiological signals are temporal series containing a lot of information, and their analysis (either for diagnosis or evolution monitoring) necessitates tools that take into account their intrinsic characteristics, notably in terms of unpredictability and high number of parameters. Methodologies coming from chaotic and nonlinear dynamical systems contain some useful building blocks in that perspective, and allow a qualitative link with phenomenological and bio-inspired models. The objective of this work is to introduce some methods in nonlinear dynamics useful for the processing of these types of signals. An application of these tools is

illustrated in the processing of potential electrical fields acquired from in vitro culture cells on newborn rats. Both normal (regular contraction of cells) and arrhythmic (disordered contractions) cases are contemplated.

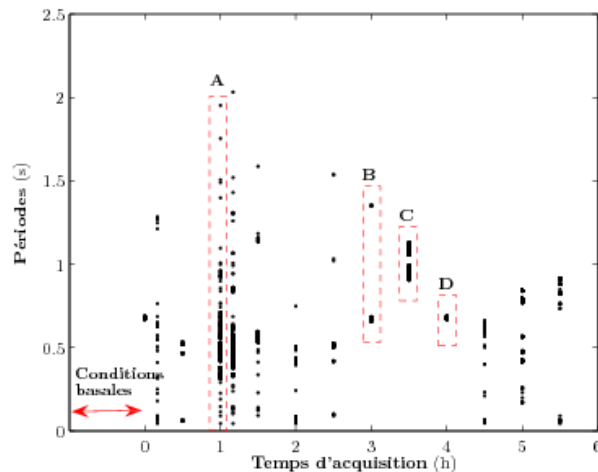


Figure 2. CPE period bifurcation diagram. Cells are stimulated by M1 electrode during 5 minutes with an impulsion train of $200\mu\text{m}$ and frequency 100Hz. Three particular phenomena in cell behaviour: **A** ($t = 1$ hour) chaotic state, **B** ($t = 3$ hour) and **C** ($t = 3.5$ hour) period doubling phase, **D** ($t = 4$ hour) regular and stable rhythm.

The bifurcation diagram is an example of a tool that can be used in the temporal analysis of an experimental system.

Publication: [32].

7.5. Nonlinear trend removal should be carefully performed in heart rate variability analysis

Participants: Binbin Xu, Oriol Pont, Hussein Yahia, Rémi Dubois.

Background : In Heart rate variability analysis, the rate-rate time series suffer often from aperiodic non-stationarity, presence of ectopic beats etc. It would be hard to extract helpful information from the original signals. **Problem :** Trend removal methods are commonly practiced to reduce the influence of the low frequency and aperiodic non-stationary in RR data. This can unfortunately affect the signal and make the analysis on detrended data less appropriate. **Objective :** Investigate the detrending effect (linear & nonlinear) in temporal / nonlinear analysis of heart rate variability of long-term RR data (in normal sinus rhythm, atrial fibrillation, congestive heart failure and ventricular premature arrhythmia conditions). **Methods :** Temporal method : standard measure SDNN; Nonlinear methods : multi-scale Fractal Dimension (FD), Detrended Fluctuation Analysis (DFA) & Sample Entropy (SampEn) analysis.

Results : The linear detrending affects little the global characteristics of the RR data, either in temporal analysis or in nonlinear complexity analysis. After linear detrending, the SDNNs are just slightly shifted and all distributions are well preserved. The cross-scale complexity remained almost the same as the ones for original RR data or correlated. Nonlinear detrending changed not only the SDNNs distribution, but also the order among different types of RR data. After this processing, the SDNN became indistinguishable between SDNN for normal sinus rhythm and ventricular premature beats. Different RR data has different complexity signature. Nonlinear detrending made the all RR data to be similar, in terms of complexity. It is thus impossible

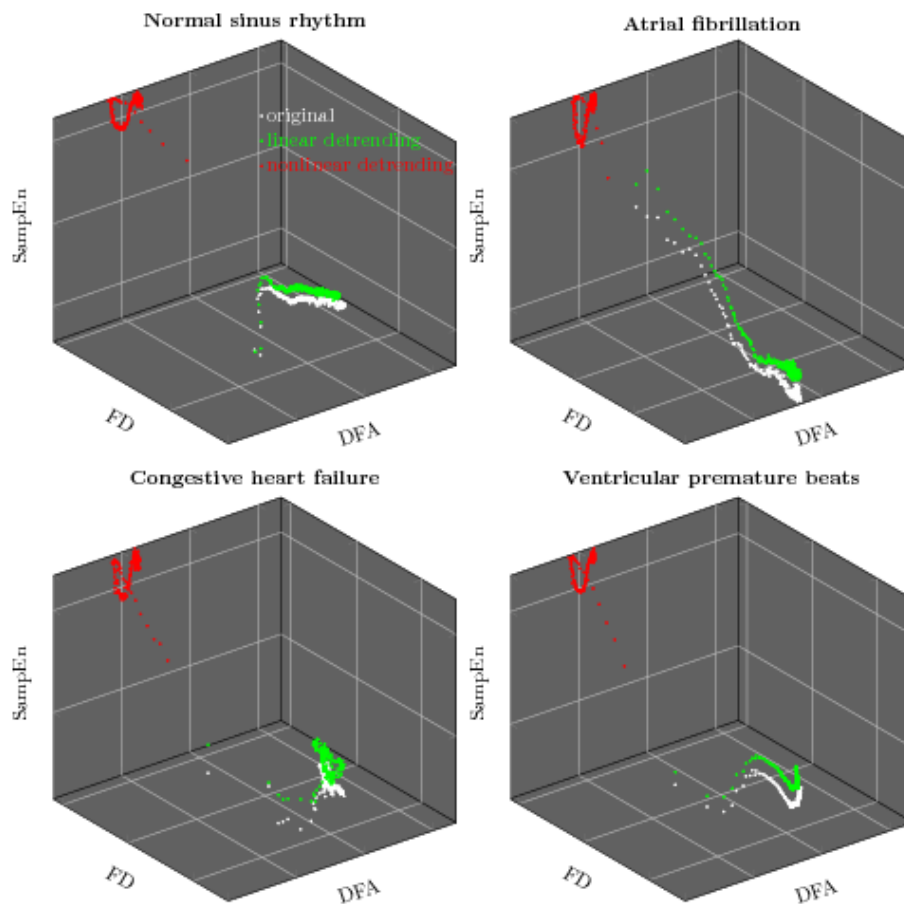


Figure 3. Complexity Space based on FD, DFA & SampEn, for RR data in NSR, AF, CHF & VPB conditions.

to distinguish them. The FD showed that nonlinearly detrended RR data has a dimension close to 2, the exponent from DFA is close to zero and SampEn is larger than 1.5 – these complexity values are very close to those for random signal. **Conclusions** : Pre-processing by linear detrending can be performed on RR data, which has little influence on the corresponding analysis. Nonlinear detrending could be harmful and it is not advisable to use this type of pre-processing. Exceptions do exist, but only combined with other appropriate techniques to avoid complete change of the signal's intrinsic dynamics. One submitted publication.

7.6. Quantification of Heart's Recover by Multiscale Complexity Analysis of Heart Rate: a Validation Study

Participants: Binbin Xu, Hussein Yahia, Rémi Dubois.

Background : Heart rate analysis is the common analysis of heart's function. After the drug treatment of cardiac arrhythmia, the heart rate looks like the same as the group with normal sinus rhythm.

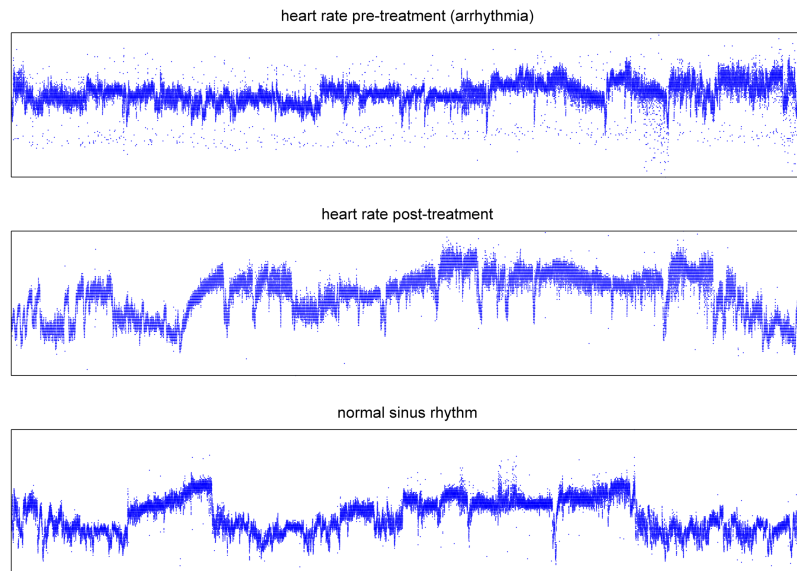


Figure 4. Heart rate RR time series in three cases : (1) arrhythmia, pre-treatment; (2) arrhythmia suppressed, post-treatment; (3) normal healthy group.

Problem : However, the visibly "same dynamics" for post-treatment & normal group does not reflect the true intrinsic dynamics of the heart. **Methods** : Using multi-scale complexity analysis to quantify and qualify the heart rate's dynamics.

Results : Thought the analysis shown in time domain that the dynamics of post-treatment and normal group looked similar. Their dynamics is completely different : (1) for normal heart rate, the multiscale fractal dimension is almost linearly decreased – invariance; (2) for arrhythmic heart rate before and after treatment, they converged to a certain value. All these suggested that item after the drug treatments, the heart's function is not still fully restored and more recovering time is needed. The multiscale complexity analysis can be used to quantify the heart function's recovering and optimize the post-treatment.

One submitted publication.

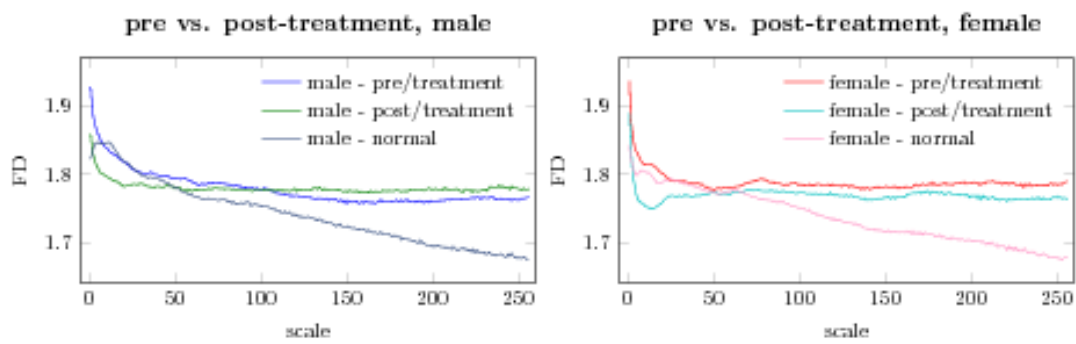


Figure 5. Analysis of heart rate variation by multiscale (coarse-graining) fractal dimension

7.7. Quantification and Its Approximate Solution of Action Potential in Neuron Models by Anharmonicity Analysis

Participants: Binbin Xu, Hussein Yahia, Rémi Dubois.

Action potential (AP) plays an important role to initiate and maintain the cell-cell communication. The nerve impulses are extensively studied but the action potential is less investigated as in other types of cells (for example, cardiac action potential). The AP can tell more about the state of the cell. It reflects the physical / chemical intracellular exchanges. Any changes in the cell would change the form/geometry of AP, or a more relevant term the *harmony*. The intrinsic changes would modify the harmony of the impulses train. The broken harmony (form/geometry change) of the impulse train means that there would be some problems in the cell. This provides an indirect way to study the intrinsic dynamics the cells.

In the work of P. Hanusse proposed a very interesting signal analyzing approach by anharmonicity, especially for signals with nonlinear oscillations properties exhibited in many physical / biological systems. This is exactly the case for neuron impulses trains. The principle is to describing the signal with their harmonic behaviors by solving the nonlinear phase equation. The obtained phase is thus used to reconstruct a solution of the original signal. The key notion is the nonlinear trigonometry that they developed. According to this approach, for any periodic signal $x(t) = x_0 + x_1 \cos(\phi(t))$, its phase can be obtained by the proposed general solution $t(\phi) = \phi + \sum_{k=1}^n a_k \text{hpsin}_1(\phi - p_k, r_k) - b_k \text{hpcos}_1(\phi - p_k, r_k)$ which can be used to reconstruct the original signal $x(t)$.

There is no practical implementation in their papers. Here we propose a first order solution of the original analytical equation. It can be used to quantify the harmonicity of the action potential.

$$\text{hpsin}_1 = \frac{-i(-\ln(1 - e^{it}r) + \ln(1 - re^{-it}))}{2r}, \quad \text{hpcos}_1 = -\frac{\ln(1 - e^{it}r) + \ln(1 - re^{-it})}{2r}$$

The phase of a signal can be solved as : $\phi(t) = t - t_0 + a_1 \text{hpsin}_1(t - t_1, r) - a_2 \text{hpsin}_1(t - t_2, r)$, so the signal can be reconstructed as $x(t) = \cos(\phi(t))$. The related parameters can be obtained by regression or nonlinear optimization methods. In consequence, all AP can be quantified by the anharmonicity parameter r .

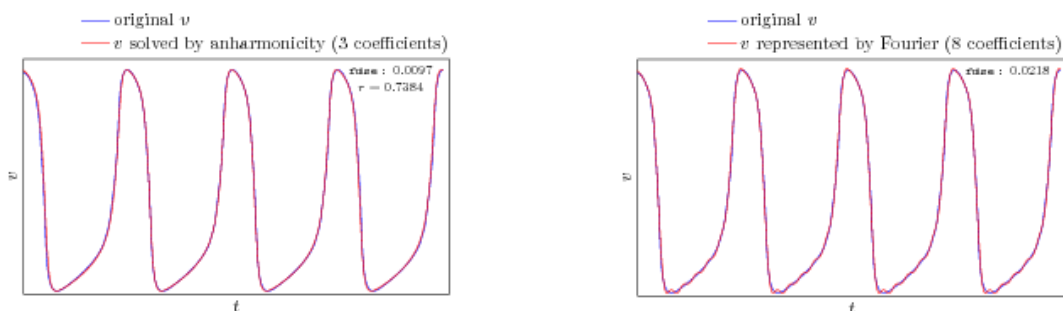


Figure 6. Quantification of action potential in FitzHugh-Nagumo (FHN) model by anharmonicity analysis

FitzHugh-Nagumo (FHN) model has been one of the basic models to study the action potential's dynamics. It's derived from the Hodgkin-Huxley model and is physiologically correct. We take here dynamics of APs from FHN model to illustrate the anharmonicity analysis. As shown in the following figure, anharmonicity analysis is far more efficient than Fourier representation. Even with 8 Fourier terms, the signal is ill-represented with Gibbs phenomena. In contrast, 3 anharmonic terms exhibit a quasi-identical fit. These AP are quantified with an anharmonicity $r = 0.7384$.

More development of anharmonicity analysis on AP is ongoing in order to provide a more efficient way to generalize the practices and for better solutions. We believe that anharmonicity analysis can help to quantify/qualify the AP in a different yet efficient way than conventional analysis.

One submitted publication.

7.8. Image Reconstruction from Highly Corrupted Gradients

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

Surface-from-Gradients (SfG) is an important step in many imaging applications. It consists in reconstructing an image/surface from corrupted gradient fields, which results in an ill-posed problem. We propose to use sparsity to regularize the problem. The first approach uses sparsity in the gradient field together with a robust norm on the data-fitting term and was presented at CVPR 2014. The new approach uses a non-local regularization that manipulates non-local similar patches of the corrupted gradient and forcing them to be low-rank. The two approaches significantly outperform previous optimization-based SfG methods on both synthetic and real data.

One submitted publication.

7.9. Fast Image Edge-Aware Processing

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

We present a framework for fast edge-aware processing of images and videos. This is an extension of our previous SIGGRAPH Asia 2013 paper. The proposed approach uses non-convex sparsity on the gradients of the latent smooth image to better preserve sharp edges. We develop tools based on first order proximal estimation for fast processing. We also propose fast and efficient numerical solutions based on separable filters estimation, which enables our method to perform fast high-quality smoothing on large-scale images. Extensive experiments show that the proposed method produces high-quality smoothing compared to state-of-the-art methods, while being fast and simple to implement.

Publication: [15].

7.10. Low-Rankness Transfer for Realistic Denoising

Participants: Hicham Badri, Hussein Yahia.

Image restoration is a very challenging task in low-level vision and is extensively used in many imaging applications. Sparsity in various forms (dictionary learning, low-rank estimation,...) has shown to be the key for successful image denoising. However, the standard noise model used to validate the results is mainly Gaussian and uniform, with known standard deviation. Unfortunately, these assumptions do not hold for real camera noise. Instead of using sparsity to model the singular values of non-local clean similar patches, we use a learning model that trains a mapping between the noisy and ground-truth clean singular values. The training is performed on real camera noise, contrary to previous methods. Experiments show that the proposed method significantly outperforms previous denoising works on real non-uniform noise and does not require estimating the standard deviation of the corruption.

One publication accepted with minor revision at IEEE Transactions on Image Processing, publication date: 2016.

7.11. Turbulent Flow Estimation

Participants: Hicham Badri, Hussein Yahia.

We use singularity exponents (SE) to regularize the problem of turbulent flow estimation under the assumption that the brightness constancy constraint holds also for (SE). We also use weighted filtering (Lucas-Kanade's solution) and sparsity on the data-fitting term to improve robustness to outliers. The proposed motion estimation is built on a Gaussian pyramid and uses the theory of warping for a better estimation of large displacements. Experiments on synthetic data show that the proposed method outperforms sophisticated methods while being simple.

Work in progress.

7.12. Pathological voice classification

Participants: Khalid Daoudi, Nicolas Brodu.

Based on our GCI detection algorithm, we redefined the classical pitch perturbation measures that are widely used in voice quality assessment. We showed that our perturbation measures yield significantly better performance in pathological voice classification than classical measures. We also showed that some matching pursuit features can allow good performances in discrimination between pathological voice categories.

Publications: [31], [30].

7.13. Emotion detection: project with Batvoice start-up

Participants: Khalid Daoudi, Nicolas Brodu.

Geostat has been granted in 2015 a Carnot-Inria contact to fund a 1 year engineer to develop a prototype of a speech emotion detection system.

7.14. Heartbeat signal analysis: Proof of Concept with IHU LIRYC

Participants: Hussein Yahia, Guillaume Attuel, Oriol Pont, Binbin Xu.

Geostat has been granted in 2015 a fund from Inria DGT to conduct pre-clinical validation from patient database acquired by IHU LIRYC.

I4S Project-Team

7. New Results

7.1. Reflectometry

7.1.1. *Experimental validation of the inverse scattering method for distributed characteristic impedance estimation*

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Florent Loete (GEEPS-SUPELEC) and with Michel Sorine, formerly member of the Inria SISYPHE EPI.

Recently published theoretic results and numerical simulations have shown the ability of inverse scattering-based methods to diagnose soft faults in electric cables, in particular, faults implying smooth spatial variations of cable characteristic parameters. The purpose of the present work is to realize laboratory experiments confirming the ability of the inverse scattering method for retrieving spatially distributed characteristic impedance from reflectometry measurements. Various smooth or stepped spatial variations of characteristic impedance profiles are tested. This study has been accomplished in the framework of the ANR SODDA project and the results have been published in IEEE Transactions on Antennas and Propagation [16].

7.2. Automatic control

7.2.1. *Observability conservation by output feedback and observability Gramian bounds*

Participants: Qinghua Zhang, Liangquan Zhang.

Though it is a trivial fact that the observability of a linear state space system is conserved by output feedback, it requires a rigorous proof to generalize this result to uniform complete observability, which is defined with the observability Gramian. The purpose of this work is to complete such a proof. Some issues in existing results are also discussed. The uniform complete observability of closed loop systems is useful for the analysis of some adaptive systems and of the Kalman filter. This study has been accomplished in the framework of the ITEA MODRIO project and the results have been published in Automatica [20].

7.2.2. *Weighted principal component analysis for Wiener system Identification: regularization and non-Gaussian excitations*

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Vincent Laurain (CRAN/CNRS/Université de Lorraine) and with Jiandong Wang (Peking University).

Finite impulse response (FIR) Wiener systems driven by Gaussian inputs can be efficiently identified by a well-known correlation-based method, except those involving even static nonlinearities. To overcome this deficiency, another method based on weighted principal component analysis (wPCA) has been recently proposed. Like the correlation-based method, the wPCA is designed to estimate the linear dynamic subsystem of a Wiener system without assuming any parametric form of the nonlinearity. To enlarge the applicability of this method, it is shown in this work that high order FIR approximation of IIR Wiener systems can be efficiently estimated by controlling the variance of parameter estimates with regularization techniques. The case of non-Gaussian inputs is also studied by means of importance sampling. The results of this study have been presented in [22].

7.2.3. *LPV system common state basis estimation from independent local LTI models*

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Lennart Ljung (Linköping University).

For the identification of a linear parameter varying (LPV) system steered by a scheduling variable evolving within a finite set, the local approach consists in separately estimating local linear time invariant (LTI) models corresponding to fixed values of the scheduling variable. It is shown in this work that, without any global structural assumption of the considered LPV system, the local state-space LTI models do not contain the necessary information about the similarity transformations making them coherent. Nevertheless, it is possible to estimate these similarity transformations from input-output data under appropriate input excitation conditions. These estimations result in a common state basis of the transformed local LTI models, so that they form a coherent global LPV model, suitable for numerical simulations in the case of fast scheduling variable evolutions. This study has been accomplished in the framework of the ITEA MODRIO project and the results have been presented in [39].

7.3. Damage detection and linear state analysis

7.3.1. *Vibration monitoring by eigenstructure change detection based on perturbation analysis*

Participants: Michael Doehler, Qinghua Zhang, Laurent Mevel.

Vibration monitoring, notably in the fields of civil, mechanical and aeronautical engineering, aims at detecting damages at an early stage, in general by using output-only vibration measurements under ambient excitation. In this work, a new method is developed for the detection of small changes in the eigenstructure of such systems. The main idea is to transform the multiplicative eigenstructure change detection problem to an additive one, by means of perturbation analysis based on the assumption of small eigenstructure changes. Another transformation then further simplifies the detection problem into the framework of a linear regression subject to additive white Gaussian noises, leading to a numerically efficient solution of the considered problem. Compared to existing methods, it has the advantages of focusing on chosen system parameters and efficiently addressing random uncertainties. The results of this study have been presented in [31].

7.3.2. *Stochastic hybrid system actuator fault diagnosis by adaptive estimation*

Participant: Qinghua Zhang.

Based on the interacting multiple model (IMM) estimator for hybrid system state estimation and on the adaptive Kalman filter for time varying system joint state-parameter estimation, a new algorithm, the adaptive IMM estimator, is developed in this work for actuator fault diagnosis in stochastic hybrid systems. The working modes of the considered hybrid systems are described by stochastic state-space models, and the mode transitions are characterized by a Markov model. Actuator faults are modeled as parameter changes, and the related fault diagnosis problem is solved by the proposed adaptive IMM estimator through joint state-parameter estimation. This study has been accomplished in the framework of the ITEA MODRIO project and the results have been presented in [40].

7.3.3. *Damage detection on real structures*

Participants: Dominique Siegert, Laurent Mevel.

This article presents the feasibility study of a new structure for a 10-m-span bridge deck, taking into account the possibilities offered by new and high-strength materials and the advantages of a traditional environmental-friendly material. Small localized damages are hardly detected by global monitoring methods. The effectiveness of vibration-based detection depends on the accuracy of the modal parameter estimates and is limited by the low sensitivity of the modal parameters to a local stiffness reduction. This paper presents the application of SSDD to detect the change of the modal parameters of the investigated structure. Further analysis with a finite element model was conducted for assessing the consistency of the expected location and extent of the damaged elements. [15].

7.3.4. *Damage detection and simulated validation*

Participants: Michael Doehler, Laurent Mevel, Saeid Allahdadian.

This section is devoted to the numerical and theoretical validation of stochastic subspace damage detection. Sample length and sensor noise robustness were investigated. [24], [23], [25].

7.3.5. *Damage quantification*

Participants: Michael Doehler, Laurent Mevel.

Fault detection for structural health monitoring has been a topic of much research during the last decade. Localization and quantification of damages, which are linked to fault isolation, have proven to be more challenging, and at the same time of higher practical impact. While damage detection can be essentially handled as a data-driven approach, localization and quantification require a strong connection between data analysis and physical models. This paper builds upon a hypothesis test that checks if the mean of a Gaussian residual vector whose parameterization is linked to possible damage locations has become non-zero in the faulty state. It is shown how the damage location and extent can be inferred and robust numerical schemes for their estimation are derived based on QR decompositions and minmax approaches. Finally, the relevance of the approach is assessed in numerical simulations of two structures.[30].

7.3.6. *Optical fiber for damage detection*

Participant: Dominique Siegert.

A technique has been developed to detect and quantify structural damages. It consists of updating the model parameters associated to the damage, i.e. Young modulus, from strain sensor outputs obtained by optical fiber. Early damage detection can be expected using the local information given by the strain measurement. The method has been applied to a 8 meter post-tensioned concrete beam under a static loading. The model updating problem can be formulated as a minimization problem, i.e minimize a data misfit functional. To solve this problem, we use a gradient-based method. The gradient of the functional is computed at a low computational cost by means of the adjoint state. The technique is able to detect the damaged area in a post-tensioned concrete beam and to estimate its level of damage. [38]

7.4. Smarts roads and R5G

7.4.1. *Positive surface temperature pavement*

Participants: Jean Dumoulin, Nicolas Le Touz.

The mobility during winter season in France mainly relies on the use of de-icers, with an amount ranging from two hundreds thousands tons up to two millions tons for the roads only. Besides the economic impact, there are many concerns on their environmental and infrastructure, both on roads and on airports. In such context and in the framework of the R5G (5th Generation Road) project driven by IFSTTAR, investigations were carried out on the way to modify the infrastructure to maintain pavement surface at a temperature above water freezing point. Two distinct approaches, that can could be combined, were selected. The first one consisted in having a heated fluid circulating in a porous layer within an asphalt concrete pavement sample. The second one specifically relied on the use of paraffin phase change materials (PCM) in cement concrete pavement ones. Experiments on enhanced pavement samples were conducted in a climatic chamber to simulate winter conditions for several continuous days, including wind and precipitations, and monitored by infrared thermography. [45], [34]

7.4.2. *Road structure design with energy harvesting capabilities*

Participants: Nicolas Le Touz, Jean Dumoulin.

Facing the heavy organisational, financial and environmental constraints imposed by usual winter maintenance salting operations, pavement engineers have been led to look for alternative solutions to avoid ice or snow deposit at pavements surface. Among the solutions, one is self-de-icing heating pavements, for which two technologies have been developed so far: one is based on embedded coils circulating a heated calorific fluid under the pavement surface; the other one relies on the use of embedded resistant electric wires. The use and operation of such systems in the world is still limited and was only confined to small road stretches or specific applications, such as bridges which are particularly sensitive to frost. One of the most significant coil technology example in Europe is the SERSO-System (Solar Energy recovery from road surfaces) built in 1994, on a Switzerland bridge. Many of these experiences are referenced in the technical literature, which provides state-of-the art papers (see for instance Eugster) and useful detailed information dealing with the construction and operational management of such installation. The present study is taking part of the Forever Open Road Concept addressed by the R5G: 5th Generation Road, one of the major project supported by IFSTTAR. It considers a different design of self-de-icing road that simplify its mode of construction and maintenance, compared to the two technologies mentioned above. It should also be noted that similar to pavements instrumented with coils, such structure could be used in the reversible way to capture the solar energy at the pavement surface during sunny days and store it, to either warm the pavement at a later stage or for exogenous needs (e.g. contribution to domestic hot water). To complete our study we also considered the use of semi-transparent pavement course wearing in place of the traditional opaque one. In the present study, a 2D model was developed using FEM approach. It combines 2 numerical models. One is dedicated to the calculation of the heat transfer inside the porous layer between the fluid and the structure according to the geometry studied and the physical properties of the components of the system. The second one addresses the heat transfer inside the different layer of the pavement and was adapted to allow the insertion of a semi-transparent surface layer (for sun radiation). The temperature spatial distribution within the structure and its surface is calculated at different time step according to the evolution of boundary conditions at its surface. Various location in France were selected and calculation of the temperature field was carried-out over a year. Discussion on the performances of such system versus its location is proposed. Influence of a semi-transparent layer is also discussed. Future works will compared numerical simulations with experiments thank to a dedicated test bench under development and that will allow to test various structure in parallel. [32]

7.5. Non Destructive Testing using Infrared Thermography

7.5.1. Optimal designs of experience for thermal NDT

Participants: Antoine Crinière, Jean Dumoulin.

During previous works, square pulsed thermography was used to carry out non destructive testing of bonding quality of CFRP glued on civil engineering structures during reinforcement operations. The use of such wave form excitation was motivated by “on-site” requirements, but also by measurements duration, number of composite layers to test, depth of possible faulting areas versus temperature elevation allowed at composite level according to inner heat diffusion. Nevertheless, square pulsed excitation implies to choose an adapted heat duration. This duration is directly linked to the reliability of the parameter estimator. According to these observations, an indicator able to predict the sufficient heating time when the reliability of the parameter estimator reached an asymptotic evolution behaviour was studied. Based on the absolute thermal contrast, the proposed indicator I_{ph} is defined with the maximum thermal contrast and the time delay between the heating time and the appearance of the maximum contrast. This indicator allows to take into account the detectability as well as the induced flaw temporal effect on the thermal contrast shape evolution. This paper will present the establishment of this indicator for optimal square heating time and present an analysis of results obtained with numerical simulations and laboratory experiments. [28]

7.5.2. Thermal NDT and signal processing

Participant: Jean Dumoulin.

This work deals with the detection of non-emergent small structures like mosaic, hidden under a plaster layer, with various spatial layout and nature. Three post processing approach by PPT, SVD and Polynomial analysis were conducted on this experimental and simulated data set. Results obtained are analysed and discussed. Finally, influence of IR camera used will be also addressed and discussed in the dissertation. [35]

7.6. Outdoor InfraRed Thermography

7.6.1. Vision enhancement through Infrared imaging for transport infrastructures

Participant: Jean Dumoulin.

Fog conditions are the cause of severe car accidents in European western countries because of the poor induced visibility. Its occurrence and intensity are still very difficult to forecast for weather services. Furthermore, visibility determination relies on expensive instruments and does not ease their dissemination. Lately, it has been demonstrated the benefit of infrared cameras to detect and to identify objects in fog while visibility is too low for eye detection. Over the past years, such cameras have become more cost effective. A research program between IFSTTAR and Cerema studied the possibility to retrieve visibility distance in a fog tunnel during its natural dissipation. The purpose of this work is to retrieve atmospheric visibility with a technique based on the combined use of infrared thermography, Principal Components Analysis (PCA) and Partial Least-Square (PLS) regression applied to infrared images.[44] and [17]

7.6.2. Outdoor thermal monitoring of large scale structures by infrared thermography

Participants: Jean Dumoulin, Antoine Crinière.

With the constant increase of the road traffic coupled with the ageing of transport infrastructure, studying and developing robust system which allows to monitor and assess those structures is of growing interest. Among the techniques used [1], thermal monitoring with infrared thermography appears to be a good compromise between a non-intrusive method and possible added value after post-processing of acquired data. Through the past decade studies have shown the ability to monitor concrete and asphalt structure by active IR thermography. On site measurement using passive thermography have also been studied, by applying qualitative methods and quantitative one. These methods have been used to perform punctual control of various duration (few hours to few days). However, infrared thermography, when it is used in a quantitative mode (not in laboratory conditions) and not in a qualitative mode (vision applied to survey), needs to process thermal radiative corrections on the raw data acquired in real time, to take into account the influences of the natural environment's evolution with time. The ICT system called "IrLaW" is based on a multi sensing approach. It connects and synchronizes information acquired by a weather station, a GPS and an infrared camera. To fulfill ICT objectives (OGCcompliant), a specific hardware architecture was also designed and studied to allow the whole system integration in a TCP/IP network. [29]

INOCS Team (section vide)

IPSO Project-Team

5. New Results

5.1. Uniformly accurate numerical schemes for highly-oscillatory Klein-Gordon and nonlinear Schrödinger equation

The work [20] is devoted to the numerical simulation of nonlinear Schrödinger and Klein-Gordon equations. We present a general strategy to construct numerical schemes which are uniformly accurate with respect to the oscillation frequency. This is a stronger feature than the usual so called "Asymptotic preserving" property, the last being also satisfied by our scheme in the highly oscillatory limit. Our strategy enables to simulate the oscillatory problem without using any mesh or time step refinement, and the orders of our schemes are preserved uniformly in all regimes. In other words, since our numerical method is not based on the derivation and the simulation of asymptotic models, it works in the regime where the solution does not oscillate rapidly, in the highly oscillatory limit regime, and in the intermediate regime with the same order of accuracy. The method is based on two main ingredients. First, we embed our problem in a suitable "two-scale" reformulation with the introduction of an additional variable. Then a link is made with classical strategies based on Chapman-Enskog expansions in kinetic theory despite the dispersive context of the targeted equations, allowing to separate the fast time scale from the slow one. Uniformly accurate (UA) schemes are eventually derived from this new formulation and their properties and performances are assessed both theoretically and numerically.

5.2. Higher-order averaging, formal series and numerical integration III: error bounds

In earlier works, it has been shown how formal series like those used nowadays to investigate the properties of numerical integrators may be used to construct high- order averaged systems or formal first integrals of Hamiltonian problems. With the new approach the averaged system (or the formal first integral) may be written down immediately in terms of (i) suitable basis functions and (ii) scalar coefficients that are computed via simple recursions. In [21], we show how the coefficients/basis functions approach may be used advantageously to derive exponentially small error bounds for averaged systems and approximate first integrals.

5.3. Stroboscopic averaging for the nonlinear Schrödinger equation

In [18], we are concerned with an averaging procedure, -namely Stroboscopic averaging-, for highly-oscillatory evolution equations posed in a (possibly infinite dimensional) Banach space, typically partial differential equations (PDEs) in a high-frequency regime where only one frequency is present. We construct a high-order averaged system whose solution remains exponentially close to the exact one over long time intervals, possesses the same geometric properties (structure, invariants,...) as compared to the original system, and is non-oscillatory. We then apply our results to the nonlinear Schrödinger equation on the d -dimensional torus T^d , or in R^d with a harmonic oscillator, for which we obtain a hierarchy of Hamiltonian averaged models. Our results are illustrated numerically on several examples borrowed from the recent literature.

5.4. Uniformly accurate time-splitting schemes for NLS in the semiclassical limit

In [42], we construct new numerical methods for the nonlinear Schrödinger equation in the semiclassical limit. We introduce time-splitting schemes for a phase-amplitude reformulation of the equation where the dimensionless Planck constant is not a singular parameter anymore. Our methods have an accuracy which is spectral in space, of second or fourth-order in time, and independent of the Planck constant before the formation of caustics. The scheme of second-order preserves exactly the L^2 norm of the solution, as the flow of the nonlinear Schrödinger equation does. In passing, we introduce a new time-splitting method for the eikonal equation, whose precision is spectral in space and of second or fourth-order in time.

5.5. Gyroaverage operator for a polar mesh

In [33], we are concerned with numerical approximation of the gyroaverage operators arising in plasma physics to take into account the effects of the finite Larmor radius corrections. This work extended a previous approach to polar geometries. A direct method is proposed in the space configuration which consists in integrating on the gyrocircles using interpolation operator (Hermite or cubic splines). Numerical comparisons with a standard method based on a Padé approximation are performed: (i) with analytical solutions, (ii) considering the 4D drift-kinetic model with one Larmor radius and (iii) on the classical linear DIII-D benchmark case. In particular, we show that in the context of a drift-kinetic simulation, the proposed method has similar computational cost as the standard method and its precision is independent of the radius.

5.6. Asymptotic Preserving scheme for a kinetic model describing incompressible fluids

The kinetic theory of fluid turbulence modeling developed by Degond and Lemou "Turbulence models for incompressible fluids derived from kinetic theory" (J. Math. Fluid Mech. 2002) is considered for further study, analysis and simulation. Starting with the Boltzmann like equation representation for turbulence modeling, a relaxation type collision term is introduced for isotropic turbulence. In order to describe some important turbulence phenomenology, the relaxation time incorporates a dependency on the turbulent microscopic energy and this makes difficult the construction of efficient numerical methods. To investigate this problem, we focus here on a multi-dimensional prototype model and first propose an appropriate change of frame that makes the numerical study simpler. Then, a numerical strategy to tackle the stiff relaxation source term is introduced in the spirit of Asymptotic Preserving Schemes. Numerical tests are performed in a one-dimensional framework on the basis of the developed strategy to confirm its efficiency.

5.7. Numerical schemes for kinetic equations in the diffusion and anomalous diffusion limits. Part I: the case of heavy-tailed equilibrium

In [44], we propose some numerical schemes for linear kinetic equations in the diffusion and anomalous diffusion limit. When the equilibrium distribution function is a Maxwellian distribution, it is well known that for an appropriate time scale, the small mean free path limit gives rise to a diffusion type equation. However, when a heavy-tailed distribution is considered, another time scale is required and the small mean free path limit leads to a fractional anomalous diffusion equation. Our aim is to develop numerical schemes for the original kinetic model which works for the different regimes, without being restricted by stability conditions of standard explicit time integrators. First, we propose some numerical schemes for the diffusion asymptotics; then, their extension to the anomalous diffusion limit is studied. In this case, it is crucial to capture the effect of the large velocities of the heavy-tailed equilibrium, so that some important transformations of the schemes derived for the diffusion asymptotics are needed. As a result, we obtain numerical schemes which enjoy the Asymptotic Preserving property in the anomalous diffusion limit, that is: they do not suffer from the restriction on the time step and they degenerate towards the fractional diffusion limit when the mean free path goes to zero. We also numerically investigate the uniform accuracy and construct a class of numerical schemes satisfying this property. Finally, the efficiency of the different numerical schemes is shown through numerical experiments.

5.8. Comparison of numerical solvers for anisotropic diffusion equations arising in plasma physics

In [25], we are concentrated to the comparison of numerical schemes to approximate anisotropic diffusion problems arising in tokamak plasma physics. We focus on the spatial approximation by using finite volume method and on the time discretization. This latter point is delicate since the use of explicit integrators leads to a severe restriction on the time step. Then, implicit and semi-implicit schemes are coupled to finite volumes space discretization and are compared for some classical problems relevant for magnetically confined plasmas.

It appears that the semi-implicit approaches (using ARK methods or directional splitting) turn out to be the most efficient on the numerical results, especially when nonlinear problems are studied on refined meshes, using high order methods in space.

5.9. Hamiltonian splitting for the Vlasov-Maxwell equations

In [23], a new splitting is proposed for solving the Vlasov-Maxwell system. This splitting is based on a decomposition of the Hamiltonian of the Vlasov-Maxwell system and allows for the construction of arbitrary high order methods by composition (independent of the specific deterministic method used for the discretization of the phase space). Moreover, we show that for a spectral method in space this scheme satisfies Poisson's equation without explicitly solving it. Finally, we present some examples in the context of the time evolution of an electromagnetic plasma instability which emphasizes the excellent behavior of the new splitting compared to methods from the literature.

5.10. Multiscale numerical schemes for kinetic equations in the anomalous diffusion limit

In [24], we construct numerical schemes to solve kinetic equations with anomalous diffusion scaling. When the equilibrium is heavy-tailed or when the collision frequency degenerates for small velocities, an appropriate scaling should be made and the limit model is the so-called anomalous or fractional diffusion model. Our first scheme is based on a suitable micro-macro decomposition of the distribution function whereas our second scheme relies on a Duhamel formulation of the kinetic equation. Both are Asymptotic Preserving (AP): they are consistent with the kinetic equation for all fixed value of the scaling parameter $\varepsilon > 0$ and degenerate into a consistent scheme solving the asymptotic model when epsilon tends to 0. The second scheme enjoys the stronger property of being uniformly accurate (UA) with respect to epsilon. The usual AP schemes known for the classical diffusion limit cannot be directly applied to the context of anomalous diffusion scaling, since they are not able to capture the important effects of large and small velocities. We present numerical tests to highlight the efficiency of our schemes.

5.11. High-order Hamiltonian splitting for Vlasov-Poisson equations

In [40], we consider the Vlasov-Poisson equation in a Hamiltonian framework and derive new time splitting methods based on the decomposition of the Hamiltonian functional between the kinetic and electric energy. Assuming smoothness of the solutions, we study the order conditions of such methods. It appears that these conditions are of Runge-Kutta-Nyström type. In the one dimensional case, the order conditions can be further simplified, and efficient methods of order 6 with a reduced number of stages can be constructed. In the general case, high-order methods can also be constructed using explicit computations of commutators. Numerical results are performed and show the benefit of using high-order splitting schemes in that context. Complete and self-contained proofs of convergence results and rigorous error estimates are also given.

5.12. Asymptotic Preserving numerical schemes for multiscale parabolic problems

In [45], we consider a class of multiscale parabolic problems with diffusion coefficients oscillating in space at a possibly small scale ε . Numerical homogenization methods are popular for such problems, because they capture efficiently the asymptotic behaviour as $\varepsilon \rightarrow 0$, without using a dramatically fine spatial discretization at the scale of the fast oscillations. However, known such homogenization schemes are in general not accurate for both the highly oscillatory regime $\varepsilon \rightarrow 0$ and the non oscillatory regime $\varepsilon \rightarrow 1$. In this paper, we introduce an Asymptotic Preserving method based on an exact micro-macro decomposition of the solution which remains consistent for both regimes.

5.13. Parallelization of an advection-diffusion problem arising in edge plasma physics using hybrid MPI/OpenMP programming

In [35], we present a hybrid MPI/OpenMP parallelization strategy for an advection-diffusion problem, arising in a scientific application simulating tokamak's edge plasma physics. This problem is the hotspot of the system of equations numerically solved by the application. As this part of the code is memory-bandwidth limited, we show the benefit of a parallel approach that increases the aggregated memory bandwidth in using multiple computing nodes. In addition, we designed some algorithms to limit the additional cost, induced by the needed extra inter nodal communications. The proposed solution allows to achieve good scalings on several nodes and to observe 70% of relative efficiency on 512 cores. Also, the hybrid parallelization allows to consider larger domain sizes, unreachable on a single computing node.

5.14. Numerical schemes for kinetic equations in the anomalous diffusion limit. Part II: degenerate collision frequency

In [43], which is the continuation of [44], we propose numerical schemes for linear kinetic equation which are able to deal with the fractional diffusion limit. When the collision frequency degenerates for small velocities it is known that for an appropriate time scale, the small mean free path limit leads to an anomalous diffusion equation. From a numerical point of view, this degeneracy gives rise to an additional stiffness that must be treated in a suitable way to avoid a prohibitive computational cost. Our aim is therefore to construct a class of numerical schemes which are able to undertake these stiffness. This means that the numerical schemes are able to capture the effect of small velocities in the small mean free path limit with a fixed set of numerical parameters. Various numerical tests are performed to illustrate the efficiency of our methods in this context.

5.15. Analysis of the Monte-Carlo error in a hybrid semi-Lagrangian scheme

In [17] we consider Monte-Carlo discretizations of partial differential equations based on a combination of semi-lagrangian schemes and probabilistic representations of the solutions. The goal of this paper is twofold. First we give rigorous convergence estimates for our algorithm: In a simple setting, we show that under an anti-CFL condition on the time-step δt and on the mesh size δx and for a reasonably large number of independent realizations N , we control the Monte-Carlo error by a term of order $\mathcal{O}(\sqrt{\delta t/N})$. Then, we show various applications of the numerical method in very general situations (nonlinear, different boundary conditions, higher dimension) and numerical examples showing that the theoretical bound obtained in the simple case seems to persist in more complex situations.

5.16. Resonant time steps and instabilities in the numerical integration of Schrödinger equations

In [30], we consider the linear and non linear cubic Schrödinger equations with periodic boundary conditions, and their approximations by splitting methods. We prove that for a dense set of arbitrary small time steps, there exists numerical solutions leading to strong numerical instabilities preventing the energy conservation and regularity bounds obtained for the exact solution. We analyze rigorously these instabilities in the semi-discrete and fully discrete cases.

5.17. Collisions of almost parallel vortex filaments

In [38], we investigate the occurrence of collisions in the evolution of vortex filaments through a system introduced by Klein, Majda and Damodaran, and by Zakharov. We first establish rigorously the existence of a pair of almost parallel vortex filaments, with opposite circulation, colliding at some point in finite time. The collision mechanism is based on the one of the self-similar solutions of the model, described in our previous work. In the second part of this paper we extend this construction to the case of an arbitrary number of filaments, with polygonal symmetry, that are perturbations of a configuration of parallel vortex filaments forming a polygon, with or without its center, rotating with constant angular velocity.

5.18. On numerical Landau damping for splitting methods applied to the Vlasov-HMF model

In [49] we consider time discretizations of the Vlasov-HMF (Hamiltonian Mean-Field) equation based on splitting methods between the linear and non-linear parts. We consider solutions starting in a small Sobolev neighborhood of a spatially homogeneous state satisfying a linearized stability criterion (Penrose criterion). We prove that the numerical solutions exhibit a scattering behavior to a modified state, which implies a nonlinear Landau damping effect with polynomial rate of damping. Moreover, we prove that the modified state is close to the continuous one and provide error estimates with respect to the time stepsize.

5.19. A kinetic model for the transport of electrons in a graphen layer

In [50], a kinetic model for the transport of electrons in graphene is derived with the tools of semiclassical analysis. The underlying quantum model is a massless Dirac equation, whose eigenvalues display a conical singularity responsible for non adiabatic transitions between the two modes. Our kinetic model takes the form of two Boltzmann equations coupled by a collision operator modeling these transitions. This collision term includes a Landau-Zener transfer term and a jump operator whose presence is essential in order to ensure a good energy conservation during the transitions. We propose an algorithmic realization of the semi-group solving the kinetic model, by a particle method. In the last section, a series of numerical experiments are given in order to study the influences of the various sources of errors between the quantum and the kinetic models.

5.20. Dimension reduction for dipolar Bose-Einstein condensates in the strong interaction regime

In [39], we study dimension reduction for the three-dimensional Gross-Pitaevskii equation with a long-range and anisotropic dipole-dipole interaction modeling dipolar Bose-Einstein condensation in a strong interaction regime. The cases of disk shaped condensates (confinement from dimension three to dimension two) and cigar shaped condensates (confinement to dimension one) are analyzed. In both cases, the analysis combines averaging tools and semiclassical techniques. Asymptotic models are derived, with rates of convergence in terms of two small dimensionless parameters characterizing the strength of the confinement and the strength of the interaction between atoms.

5.21. The Interaction Picture method for solving the generalized nonlinear Schrödinger equation in optics

The "interaction picture" (IP) method studied in [13] is a very promising alternative to Split-Step methods for solving certain type of partial differential equations such as the nonlinear Schrödinger equation used in the simulation of wave propagation in optical fibers. The method exhibits interesting convergence properties and is likely to provide more accurate numerical results than cost comparable Split-Step methods such as the Symmetric Split-Step method. In this work we investigate in detail the numerical properties of the IP method and carry out a precise comparison between the IP method and the Symmetric Split-Step method.

5.22. Nonlinear stability criteria for the HMF Model

In [52], we study the nonlinear stability of a large class of inhomogeneous steady state solutions to the Hamiltonian Mean Field (HMF) model. Under a simple criterion, we prove the nonlinear stability of steady states which are decreasing functions of the microscopic energy. To achieve this task, we extend to this context the strategy based on generalized rearrangement techniques which was developed recently for the gravitational Vlasov-Poisson equation. Explicit stability inequalities are established and our analysis is able to treat non compactly supported steady states to HMF, which are physically relevant in this context but induces additional difficulties, compared to the Vlasov-Poisson system.

5.23. Dimension reduction for rotating Bose-Einstein condensates with anisotropic confinement

In [54], we consider the three-dimensional time-dependent Gross-Pitaevskii equation arising in the description of rotating Bose-Einstein condensates and study the corresponding scaling limit of strongly anisotropic confinement potentials. The resulting effective equations in one or two spatial dimensions, respectively, are rigorously obtained as special cases of an averaged three dimensional limit model. In the particular case where the rotation axis is not parallel to the strongly confining direction the resulting limiting model(s) include a negative, and thus, purely repulsive quadratic potential, which is not present in the original equation and which can be seen as an effective centrifugal force counteracting the confinement.

5.24. Dimension reduction for anisotropic Bose-Einstein condensates in the strong interaction regime

In [14], we study the problem of dimension reduction for the three dimensional Gross-Pitaevskii equation (GPE) describing a Bose-Einstein condensate confined in a strongly anisotropic harmonic trap. Since the gas is assumed to be in a strong interaction regime, we have to analyze two combined singular limits: a semi-classical limit in the transport direction and the strong partial confinement limit in the transversal direction. We prove that both limits commute together and we provide convergence rates. The by-products of this work are approximated models in reduced dimension for the GPE, with a priori estimates of the approximation errors.

5.25. Models of dark matter halos based on statistical mechanics: I. The classical King model

In [22], we consider the possibility that dark matter halos are described by the Fermi-Dirac distribution at finite temperature. This is the case if dark matter is a self-gravitating quantum gas made of massive neutrinos at statistical equilibrium. This is also the case if dark matter can be treated as a self-gravitating collisionless gas experiencing Lynden-Bell's type of violent relaxation. In order to avoid the infinite mass problem and carry out a rigorous stability analysis, we consider the fermionic King model. In this paper, we study the non-degenerate limit leading to the classical King model. This model was initially introduced to describe globular clusters. We propose to apply it also to large dark matter halos where quantum effects are negligible. We determine the caloric curve and study the thermodynamical stability of the different configurations. Equilibrium states exist only above a critical energy E_c in the microcanonical ensemble and only above a critical temperature T_c in the canonical ensemble.

5.26. Numerical study of a quantum-diffusive spin model for two-dimensional electron gases

In [15], we investigate the time evolution of spin densities in a two-dimensional electron gas subjected to Rashba spin-orbit coupling on the basis of the quantum drift-diffusive model. This model assumes the electrons to be in a quantum equilibrium state in the form of a Maxwellian operator. The resulting quantum drift-diffusion equations for spin-up and spin-down densities are coupled in a non-local manner via two spin chemical potentials (Lagrange multipliers) and via off-diagonal elements of the equilibrium spin density and spin current matrices, respectively. We present two space-time discretizations of the model, one semi-implicit and one explicit, which comprise also the Poisson equation in order to account for electron-electron interactions. In a first step pure time discretization is applied in order to prove the well-posedness of the two schemes, both of which are based on a functional formalism to treat the non-local relations between spin densities. We then use the fully space-time discrete schemes to simulate the time evolution of a Rashba electron gas confined in a bounded domain and subjected to spin-dependent external potentials. Finite difference approximations are first order in time and second order in space. The discrete functionals introduced are minimized with the help of a conjugate gradient-based algorithm, where the Newton method is applied in

order to find the respective line minima. The numerical convergence in the long-time limit of a Gaussian initial condition towards the solution of the corresponding stationary Schrödinger- Poisson problem is demonstrated for different values of the numerical parameters. Moreover, the performances of the semi-implicit and the explicit scheme are compared.

5.27. Numerical analysis of the nonlinear Schrödinger equation with white noise dispersion

In [16], we focus to the numerical study of a nonlinear Schrödinger equation in which the coefficient in front of the group velocity dispersion is multiplied by a real valued Gaussian white noise. We first perform the numerical analysis of a semi-discrete Crank-Nicolson scheme in the case when the continuous equation possesses a unique global solution. We prove that the strong order of convergence in probability is equal to one in this case. In a second step, we numerically investigate, in space dimension one, the behavior of the solutions of the equation for different power nonlinearities, corresponding to subcritical, critical or supercritical nonlinearities in the deterministic case. Numerical evidence of a change in the critical power due to the presence of the noise is pointed out.

5.28. A regularity result for quasilinear stochastic partial differential equations of parabolic type

In [27], we consider a quasilinear parabolic stochastic partial differential equation driven by a multiplicative noise and study regularity properties of its weak solution satisfying classical a priori estimates. In particular, we determine conditions on coefficients and initial data under which the weak solution is Hölder continuous in time and possesses spatial regularity that is only limited by the regularity of the given data. Our proof is based on an efficient method of increasing regularity: the solution is rewritten as the sum of two processes, one solves a linear parabolic SPDE with the same noise term as the original model problem whereas the other solves a linear parabolic PDE with random coefficients. This way, the required regularity can be achieved by repeatedly making use of known techniques for stochastic convolutions and deterministic PDEs.

5.29. Diffusion limit for the radiative transfer equation perturbed by a Wiener process

The aim of [28] is the rigorous derivation of a stochastic non-linear diffusion equation from a radiative transfer equation perturbed with a random noise. The proof of the convergence relies on a formal Hilbert expansion and the estimation of the remainder. The Hilbert expansion has to be done up to order 3 to overcome some difficulties caused by the random noise.

5.30. Invariant measure of scalar first-order conservation laws with stochastic forcing

In [29], under an hypothesis of non-degeneracy of the flux, we study the long-time behaviour of periodic scalar first-order conservation laws with stochastic forcing in any space dimension. For sub-cubic fluxes, we show the existence of an invariant measure. Moreover for sub-quadratic fluxes we show uniqueness and ergodicity of the invariant measure. Also, since this invariant measure is supported by L^p for some p small, we are led to generalize to the stochastic case the theory of L^1 solutions developed by Chen and Perthame in 2003.

5.31. An integral inequality for the invariant measure of a stochastic reaction-diffusion equation

In [46], we consider a reaction-diffusion equation perturbed by noise (not necessarily white). We prove an integral inequality for the invariant measure ν of a stochastic reaction-diffusion equation. Then we discuss some consequences as an integration by parts formula which extends to ν a basic identity of the Malliavin Calculus. Finally, we prove the existence of a surface measure for a ball and a half-space of H .

5.32. Estimate for $P_t D$ for the stochastic Burgers equation

In [47], we consider the Burgers equation on $H = L^2(0, 1)$ perturbed by white noise and the corresponding transition semigroup P_t . We prove a new formula for $P_t D\varphi$ (where $\varphi : H \rightarrow \mathbb{R}$ is bounded and Borel) which depends on φ but not on its derivative. Then we deduce some new consequences for the invariant measure ν of P_t as its Fomin differentiability and an integration by parts formula which generalises the classical one for gaussian measures.

5.33. Existence of the Fomin derivative of the invariant measure of a stochastic reaction-diffusion equation

In [48], we consider a reaction-diffusion equation perturbed by noise (not necessarily white). We prove existence of the Fomin derivative of the corresponding transition semigroup P_t . The main tool is a new estimate for $P_t D\varphi$ in terms of $\|\varphi\|_{L^2(H, \nu)}$, where ν is the invariant measure of P_t .

5.34. Global behavior of N competing species with strong diffusion: diffusion leads to exclusion

In [19], we study the following problem. For a large class of models involving several species competing for a single resource in a *homogeneous* environment, it is known that the competitive exclusion principle holds: only one species survives eventually. Various works indicate though that coexistence of many species is possible when the competition occurs in a *heterogeneous* environment. We propose here a spatially heterogeneous system modeling several species competing for a single resource, and migrating in the spatial domain. For this model, it is known, at least in particular cases, that if migrations are *slow* enough, then coexistence occurs. In this paper we show at variance that if the spatial migrations are *fast* enough, then our system can be approximated by a spatially homogeneous system, called aggregated model, which can be explicitly computed, and we show that if the competitive exclusion principle holds for the aggregated model, then it holds as well for the original, spatially heterogeneous model. In other words, we show the persistence of the competitive exclusion principle in the spatially heterogeneous situation when migrations are fast. As a consequence, for fast migrations only one species may survive, namely the best competitor *in average*. We last study which is the best competitor *in average* on some examples, and draw some ecological consequences.

5.35. Randomized message-passing test-and-set

In [37] and [34], we present a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes, returns yes to a unique process and returns no to all the others. Recently many advances in implementing Test&Set objects have been achieved, however all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p \leq n$ of processes where n is the total number of processes in the system. It has an expected individual step complexity in $O(\log p)$ against an oblivious adversary, and an expected individual message complexity in $O(n)$. The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

MATHERIALS Project-Team

6. New Results

6.1. Electronic structure calculations

Participants: Eric Cancès, Virginie Ehrlacher, David Gontier, Claude Le Bris, Gabriel Stoltz.

In electronic structure calculation as in most of our scientific endeavors, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

E. Cancès and N. Mourad have clarified the mathematical framework underlying the construction of norm-conserving semilocal pseudopotentials for Kohn-Sham models, and have proved the existence of optimal pseudopotentials for a family of optimality criteria [34].

E. Cancès and R. Scott (University of Chicago) have examined a technique of Slater and Kirkwood which provides an exact resolution of the asymptotic behavior of the van der Waals attraction between two hydrogen atoms. They have modified their technique to make the problem more tractable analytically and more easily solvable by numerical methods [35].

In [33], E. Cancès, D. Gontier and G. Stoltz analyze the GW method for finite electronic systems. This method allows to compute excited states. To understand it, a first step is to provide a mathematical framework for the usual one-body operators that appear naturally in many-body perturbation theory. It is then possible to study the GW equations which construct an approximation of the one-body Green's function, and give a rigorous mathematical formulation of these equations. With this framework, results can be established for the well-posedness of the GW_0 equations, a specific instance of the GW model. In particular, the existence of a unique solution to these equations is proved in a perturbative regime.

D. Gontier extended his last-year result on N-representability by including the characterization of representable paramagnetic currents [42]. Together with Salma Lahbabi (former student of E. Cancès, University Hassan II Casablanca, ENSEM), he proved the exponential convergence rates of the uniform sampling of the Brillouin zone for the calculation of crystalline structure properties, in linear and nonlinear settings [43].

A. Bakhta, E. Cancès and V. Ehrlacher have recently been working on the design of an efficient numerical method to solve the inverse band structure problem. The aim of this work is the following: given a set of electronic bands partially characterizing the electronic structure of a crystal, is it possible to recover the structure of a material which could achieve similar electronic properties? The main difficulty in this problem relies in the practical resolution of an associated optimization problem with numerous local optima.

E. Cancès has pursued his long-term collaboration with Y. Maday (Paris 6) on the numerical analysis of electronic structure models. Together with G. Dusson (Paris 6), B. Stamm (Paris 6), and M. Vohralík (Inria), they have designed a new postprocessing method for planewave discretizations of nonlinear Schrödinger equations, and used it to compute sharp *a posteriori* error estimators for both the discretization error and the algorithmic error (convergence threshold in the iterations on the nonlinearity). They have then extended this approach to the Kohn-Sham model. In parallel, they have derived *a posteriori* error estimates for conforming numerical approximations of the Laplace eigenvalue problem with a homogeneous Dirichlet boundary condition [32]. In particular, upper and lower bounds for the first eigenvalue are given. These bounds are guaranteed, fully computable, and converge with the optimal speed to the exact eigenvalue.

Implicit solvation models aim at computing the properties of a molecule in solution (most chemical reactions take place in the liquid phase) by replacing all the solvent molecules but the few ones strongly interacting with the solute, by an effective continuous medium accounting for long-range electrostatics. E. Cancès, Y. Maday (Paris 6), and B. Stamm (Paris 6) have recently introduced a very efficient domain decomposition method for the simulation of large molecules in the framework of the so-called COSMO implicit solvation models. In collaboration with F. Lipparini (Paris 6), B. Mennucci (Department of Chemistry, University of Pisa) and J.-P.

Picquemat (Paris 6), they have implemented this algorithm in widely used computational software products (Gaussian and Tinker). E. Cancès, Y. Maday, F. Lipparini and B. Stamm have also extended this approach to the more complex polarizable continuum model (PCM).

C. Le Bris, in collaboration with P. Rouchon (École des Mines de Paris) and with J. Roussel, in the context of an internship at École des Ponts, has pursued the study of a new efficient numerical approach, based on a model reduction technique, to simulate high dimensional Lindblad type equations at play in the modelling of open quantum systems. The added value of the most recent contribution with respect to the previous studies lies in two different aspects. First, the rank of the reduced model used as surrogate for the full model can now be dynamically adjusted, in an adaptive strategy. Second, a variance reduction approach based on the technique of control variate has been developed. The noise intrinsically present in the Monte-Carlo simulation of the underlying stochastic dynamics may indeed be reduced by using the deterministic reduced model as control variate. A publication collects these two aspects and reports on the results achieved [19].

6.2. Complex fluids

Participant: Sébastien Boyaval.

The aim of the research performed in the project-team about complex fluids is mainly focused on the mathematical modelling and numerical simulation of i) non-Newtonian rheologies, with application to geophysical fluids such as mudflows, or the solid transport in rivers, and ii) stratified flows, in particular free-surface flows, which naturally occur in the geophysical context under gravity influence.

The need for reduced models is crucial for numerical computations at the large geophysical scale. S. Boyaval has therefore pursued his research about a systematic asymptotic reduction technique for thin-layers of non-Newtonian fluids with a near hydrostatic pressure [11]. On the other hand, accurate numerical simulations (for benchmark purposes at least) require a full 3D model mainly based on Stokes-like equations, and there is a constant need for better computation methods in that field too. With a view to *condensed* high-order approximations of elliptic PDEs like the Stokes equation on generic meshes (obtained by refinement or agglomeration of a simplicial initial mesh), S. Boyaval has participated in a joint work about hybridization of a mixed-dual generic approach [8]. On the hydraulic applications side, the studies initiated at CEMRACS 2013 about a stochastic representation of fluctuations in the transport of river sediments by bed-load have been published [9].

6.3. Homogenization

Participants: Michael Bertin, Ludovic Chamoin, Virginie Ehrlacher, Thomas Hudson, Marc Josien, Claude Le Bris, Frédéric Legoll, Simon Lemaire, François Madiot, William Minvielle.

6.3.1. Deterministic non periodic systems

The homogenization of (deterministic) non periodic systems is a well known topic. Although well explored theoretically by many authors, it has been less investigated from the standpoint of numerical approaches (except in the random setting). In collaboration with X. Blanc (Paris 7) and P.-L. Lions (Collège de France), C. Le Bris has introduced a possible theory, giving rise to a numerical approach, for the simulation of multiscale nonperiodic systems. The theoretical considerations are based on earlier works by the same authors (derivation of an algebra of functions appropriate to formalize a theory of homogenization). The numerical endeavor is completely new. The theoretical results obtained to date are being collected in a series of manuscripts that will be available shortly. The publications [30] and [10] specifically address the issues related to a local perturbation of the periodic problem and the challenging, practically relevant problem of interfaces between periodic structures of different nature (the celebrated "twin boundaries" problem in materials science). Some related problems will now be addressed in the context of the PhD thesis of M. Josien.

6.3.2. Stochastic homogenization

The project-team has pursued its efforts in the field of stochastic homogenization of elliptic equations, aiming at designing numerical approaches that both are practically relevant and keep the computational workload limited.

Using the standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the *whole* space \mathbb{R}^d . This equation is therefore delicate and expensive to solve. In practice, the space \mathbb{R}^d is truncated to some bounded domain, on which the corrector problem is numerically solved. In turn, this yields a converging approximation of the homogenized tensor, which happens to be a *random* matrix.

In [47], C. Le Bris, F. Legoll and W. Minvielle have investigated the possibility to use a variance reduction technique based on computing the corrector equation *only for selected environments*. These environments are chosen based on the fact that their statistics in the finite supercell matches the statistics of the materials in the infinite supercell. This method yields an approximation of the homogenized matrix with an error smaller than standard approximations. The efficiency of the approach has been demonstrated for various types of random materials, including composite materials with randomly located inclusions.

In addition, M. Bertin and F. Legoll, in collaboration with S. Brisard (École des Ponts), have investigated the possibility to use the Hashin-Shtrikman bounds as control variables in a control variate approach. The Hashin-Shtrikman bounds are often used in the computational mechanics community as approximations of the homogenized quantities. Our aim is use them to improve the efficiency of the reference computations, somewhat in the spirit of a preconditionner. Preliminary encouraging numerical results have been obtained.

Over the past years, the project-team has proposed several variance reduction techniques, see e.g. [21] for a method using antithetic variables (in a nonlinear context) and [20] for a control variate approach using a surrogate model based on a defect-type theory. These various approaches have been reviewed and compared to one another in [29].

In collaboration with B. Stamm (Paris 6), E. Cancès, V. Ehrlacher and F. Legoll have proposed in [13] a new approach to approximate the homogenized coefficients of a random stationary material. This method is an alternative to that proposed e.g. by A. Bourgeat and A. Piatniski in [Approximations of effective coefficients in stochastic homogenization, Annales de l'Institut Henri Poincaré 40, 2004] which consists in solving a corrector problem on a bounded domain. The method introduced in [13] is based on a new corrector problem, which is posed on the entire space, but which is simpler than the standard corrector problem in that the coefficients of the equation are uniform outside some ball of finite radius. This implies that, in some cases (including the case of randomly located spherical inclusions), this new corrector problem can be recast as an integral equation posed on the surface of the inclusions. The problem can then be efficiently solved via domain decomposition and using spherical harmonics.

6.3.3. Multiscale Finite Element approaches

From a numerical point of view, the Multiscale Finite Element Method (MsFEM) is a classical strategy to address the situation when the homogenized problem is not known (e.g. in difficult nonlinear cases), or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as an accurate enough approximation).

The MsFEM has been introduced more than 10 years ago. However, even in simple deterministic cases, there are still some open questions, for instance concerning multiscale advection-diffusion equations. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interacts with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes. In that spirit, C. Le Bris, F. Legoll and F. Madiot have studied in [46] several variants of the Multiscale Finite Element Method (MsFEM), specifically designed to address multiscale advection-diffusion problems in the convection-dominated regime. Generally speaking, the idea of the MsFEM is to perform a Galerkin approximation of the problem using specific basis functions, that are precomputed (in an offline stage) and adapted to the problem considered. Several possibilities for the basis functions have been examined (for instance, they may or may not encode the convection field). Depending on how basis functions are defined, stabilization techniques (such as SUPG) may be required. Another option to handle such problems is to use a splitting approach, with two legacy codes, one solving a purely diffusive multiscale equation, the other one solving a single scale, convection-dominated advection-diffusion equation. In [46], these various approaches have been compared in terms of accuracy and computational costs.

In the context of the PhD thesis of F. Madiot, current efforts are focused on the study of an advection-diffusion equation with a dominating convection in a *perforated domain*. The multiscale character of the problem here stems from the geometry of the domain. A paramount difference with the case considered in [46] is that boundary layers may appear throughout the domain (i.e. in the neighborhood of each perforation). The accuracy of the numerical approaches in the boundary layers thus becomes critical.

Most of the numerical analysis studies of the MsFEM are focused on obtaining *a priori* error bounds. In collaboration with L. Chamoin, who is currently in delegation in the project-team for the second year (from ENS Cachan, since September 2014), members of the project-team have been working on *a posteriori* error analysis for MsFEM approaches, with the aim to develop error estimation and adaptation tools. They have extended to the MsFEM case an approach that is classical in the computational mechanics community for single scale problems, and which is based on the so-called Constitutive Relation Error (CRE). Once a numerical solution u_h has been obtained, the approach needs additional computations in order to determine a divergence-free field as close as possible to the exact flux $k\nabla u$. In the context of the MsFEM, it is important to be able to do all the expensive computations in an offline stage, independently of the right-hand side. The standard CRE approach thus needs to be adapted to that context, in order to keep that feature that makes it adapted to a multiscale, multi-query context. The proposed approach yields very interesting results, and provide an accurate and robust estimation of the global error.

Current efforts are targeted towards the design of adaptive algorithms for specific quantities of interest (in the so-called “goal-oriented” setting), and the design of model reduction approaches (such as the Proper Generalized Decomposition, or PGD) in the specific context of multiscale problems.

6.3.4. Coarse approximation of an elliptic problem with oscillatory coefficients

Still another question investigated in the project-team is to find an alternative to standard homogenization techniques when the latter are difficult to use in practice. Consider a linear elliptic equation, say in divergence form, with a highly oscillatory matrix coefficient, and assume that this problem is to be solved for a large number of right-hand sides. If the coefficient oscillations are infinitely rapid, the solution can be accurately approximated by the solution to the homogenized problem, where the homogenized coefficient has been evaluated beforehand by solving the corrector problem. If the oscillations are moderately rapid, one can think instead of MsFEM-type approaches to approximate the solution to the reference problem. However, in both cases, the complete knowledge of the oscillatory matrix coefficient is required, either to build the average model or to compute the multiscale basis. In many practical cases, this coefficient is often only partially known, or merely completely unavailable, and one only has access to the solution of the equation for some loadings. This observation has led to think about alternative methods, in the following spirit. Is it possible to approximate the reference solution by the solution to a problem with a *constant* matrix coefficient? How can this “best” constant matrix approximating the oscillatory problem be constructed in an efficient manner?

A preliminary step, following discussion and interaction with A. Cohen (Paris 6), has been to cast the problem as a convex optimization problem. We have then shown that the “best” constant matrix defined as the solution of that problem converges to the homogenized matrix in the limit of infinitely rapidly oscillatory coefficients. Furthermore, the optimization problem being convex, it can be efficiently solved using standard algorithms. C. Le Bris, F. Legoll and S. Lemaire have comprehensively explored that problem. The algorithm can be made very efficient, and it yields accurate approximation of the homogenized matrix. We have also shown that it is possible to construct, in a second stage, approximations to the correctors, in order to recover an approximation of the *gradient* of the solution.

6.3.5. Optimization of a material microstructure

A project involving V. Ehrlacher and F. Legoll, in collaboration with G. Leugering and M. Stingl (Cluster of Excellence, Erlangen-Nuremberg University), aims at optimizing the shape of some materials (modeled as structurally graded linear elastic materials) in order to achieve the best mechanical response at the minimal cost. As is often the case in shape optimization, the solution tends to be highly oscillatory, hence the need for homogenization techniques. Materials under consideration are being thought of as microstructured materials composed of steel and void and whose microstructure patterns are constructed as the macroscopic deformation

of a reference periodic microstructure. The optimal material (i.e. the best macroscopic deformation) is the deformation achieving the best mechanical response.

For a given deformation, one can first compute the mechanical response using a *homogenized model*. This is the first variant that has been followed. Model reduction techniques are then required, in order to expedite the resolution of the corrector problem needed to identify the homogenized coefficient at each loop of the optimization algorithm. In that context, a PGD-type approach has been proposed.

A second variant is to compute the mechanical response at the *microscale*, using the highly oscillatory model. Preliminary results have been obtained. Current efforts are focused towards choosing an appropriate model reduction strategy.

6.3.6. Discrete systems and their thermodynamic limit

We conclude this section by describing works of the project-team on discrete models with highly oscillatory coefficients.

Dislocations are geometric line defects which interact via long-range stress fields in crystalline solids. In [45], T. Hudson has studied the thermally-driven motion of dislocations in a discrete Monte Carlo model, showing that over long observation times at low temperature in a large body, the most probable trajectory of straight dislocation lines lie close to the solution of an explicit deterministic evolution equation.

Another work is related to the understanding of the origin of hysteresis in rubber-made materials. When submitted to cyclic deformations, the strain-stress curve of these materials indeed shows a hysteresis behavior, which seems to be independent of the speed of loading. Some years ago, members of the project-team have suggested a model, at a mesoscale, to explain this behavior. This model was written in terms of a system made of a finite number of particles. F. Legoll, T. Lelièvre and T. Hudson are currently studying whether a thermodynamic limit of the model previously proposed can be identified. In order to simplify the setting, the reference discrete model has been replaced by a continuum model with highly oscillatory coefficients. This model is nonlinear and time-dependent. The question is now to identify (e.g. using two-scale convergence arguments) its homogenized limit, first in a periodic setting, second in a stochastic setting.

6.4. Computational Statistical Physics

Participants: Giacomo Di Gesù, Thomas Hudson, Dorian Le Peutrec, Frédéric Legoll, Tony Lelièvre, Antoine Levitt, Boris Nectoux, Julien Roussel, Mathias Rousset, Gabriel Stoltz, Pierre Terrier, Pierre-André Zitt.

The work of the project-team in this area is concentrated on two new directions: the sampling of reactive trajectories (where rare events dictate the dynamics of the system), and the computation of average properties of nonequilibrium systems (which complements the more traditional field of techniques to compute free energy differences).

6.4.1. Sampling of reactive trajectories

Finding trajectories for which the system undergoes a significant change is a challenging task since the transition events are typically very rare. Several methods have been proposed in the physics and chemistry literature, and members of the project-team have undertaken their study in the past years.

A first class of techniques are the accelerated dynamics introduced by A. Voter (Los Alamos National Lab) and his collaborators. A short review on the mathematical analysis of these dynamics was written by T. Lelièvre, see [48]. In [23], T. Lelièvre and F. Nier (Paris 13) analyze the low temperature asymptotics for Quasi-Stationary Distributions in a bounded domain. The objective of this analysis is to justify mathematically the validity of hyperdynamics.

Another class of techniques to compute reactive trajectories is based on splitting techniques. After the first result obtained in [12], C.E. Bréhier, T. Lelièvre and M. Rousset pursued their analysis of the Adaptive Multilevel Splitting algorithm, which is a rare event simulation method. In [31], a generalization of the method is proposed, and it is shown how to make the estimator unbiased in a discrete-in-time setting (which is generically the setting encountered in practice). Numerical experiments illustrate the performance of the method.

6.4.2. Nonequilibrium systems and non-reversible dynamics

In [38], T. Lelièvre has studied with A. Duncan and G.A. Pavliotis nonreversible diffusion processes to sample a probability measure. It is shown that nonreversible dynamics are always better in terms of the asymptotic variance (statistical error), but the efficiency of the whole algorithm sensitively depends on the time discretization algorithm, which may induce some bias (deterministic error).

T. Lelièvre together with R. Assaraf, B. Jourdain and R. Roux, have analyzed in [27] the validity of non equilibrium molecular dynamics techniques to compute the derivative of an observable with respect to a parameter-dependent probability measure. The probability measure is defined as the stationary state of a non-reversible stochastic dynamics (in particular no analytical formula for this measure is available). Such computations are at the basis of the numerical approximation of transport coefficients in molecular dynamics.

6.4.3. Numerical analysis of simulation techniques

In [44], G. Stoltz, together with A.-A. Homman (École des Ponts) and J.-B. Maillet (CEA/DAM), present new parallelizable numerical schemes for the integration of Dissipative Particle Dynamics with Energy conservation. So far, no numerical scheme was able to correctly preserve the energy over long times and give rise to small errors on average properties for moderately small timesteps, while being straightforwardly parallelizable. Two new methods are proposed, both of them straightforwardly parallelizable, and allowing to correctly preserve the total energy of the system. The accuracy and performance of these new schemes are illustrated both on equilibrium and nonequilibrium parallel simulations.

The discretization of overdamped Langevin dynamics, through schemes such as the Euler-Maruyama method, may lead to numerical methods which are unstable when the forces are non-globally Lipschitz. One way to stabilize numerical schemes is to superimpose some acceptance/rejection rule, based on a Metropolis-Hastings criterion for instance. However, rejections perturb the dynamical consistency of the resulting numerical method with the reference dynamics. G. Stoltz and M. Fathi (Berkeley) present in [40] some modifications of the standard stabilization of discretizations of overdamped Langevin dynamics by a Metropolis-Hastings procedure, which allow to either improve the strong order of the numerical method, or to decrease the bias in the estimation of transport coefficients characterizing the effective dynamical behavior of the dynamics. The latter approach relies on modified numerical schemes together with a Barker rule for the stabilization.

A. Levitt, in collaboration with C. Ortner (University of Warwick), has worked on the numerical analysis of saddle point search, an important step in the computation of reaction rates. While the convergence theory of minimization algorithms, such as the gradient method, is well-understood and standard, no such theory exists for saddle point algorithms such as the dimer method. Their work reveals a major obstruction to convergence: for some systems, the dimer method can oscillate indefinitely. This shows that there is no Lyapunov function for the associated flow, and highlights the fundamental difference between minimization and saddle search. Further work focuses on improving the reliability and convergence speed of such methods.

6.4.4. Free energy computations

The topic of free energy computations is still a significant research area of the project-team. T. Lelièvre has co-authored a review article [14] on the adaptive biasing force (ABF) method.

In addition, two new results have been obtained on the ABF method by H. Al Rachid (École des Ponts) in collaboration with T. Lelièvre: a numerical result concerning a projected version of the ABF algorithm, which enables to reduce the variance, see [25]; and a theoretical result on the existence of a solution to the non linear Fokker Planck equation associated to the ABF process, see [49].

T. Lelièvre and G. Stoltz, together with G. Fort (Télécom Paris) and B. Jourdain (École des Ponts), have studied the Self-Healing Umbrella Sampling (SHUS) method in [16]. This method is an adaptive biasing method to compute free energies on the fly by appropriately penalizing already visited regions. The convergence of the method relies on a rewriting as a stochastic approximation method with random steps, and can therefore be seen as a variation of the Wang-Landau method.

6.4.5. Convergence of processes

D. Le Peutrec and G. Di Gesù have studied in [37] the rate of convergence to equilibrium at low temperature of a stochastic interacting large particle system which can be seen as a spatially discrete approximation of the stochastic Allen-Cahn equation on the one-dimensional torus. Upper and lower bounds for the leading term of the associated spectral gap in the small temperature regime are proven, uniformly in the system size. It is also shown that the upper bound is sharp under a suitable control of the growth of the system size by the temperature.

The article [17] by B. Jourdain (École des Ponts), T. Lelièvre and B. Miasojedow (Warsaw) on the mean-field limit for the transient phase of the random walk Metropolis algorithm in the infinite dimension limit has been published in *Annals of Applied Probability*. In this article, the authors prove that the Metropolis Hastings algorithm converges to a nonlinear stochastic differential equation in the infinite dimensional limit.

6.4.6. Force fields and modeling

In [41], G. Stoltz, together with G. Ferré (École des Ponts) and J.B. Maillet (CEA/DAM), has presented a distance between atomic configurations, which is invariant with respect to permutations of the atoms. This distance is defined through a functional representation of atomic positions. It allows to directly compare different atomic environments with an arbitrary number of particles without going through a space of reduced dimensionality (i.e. fingerprints) as an intermediate step. Moreover, this distance is naturally invariant through permutations of atoms and through global rotations. This distance provides an important building block for the construction of accurate force-fields using machine learning techniques.

E. Cancès has contributed to the development of more efficient algorithms for polarizable force field molecular dynamics, which have been implemented and successfully tested on massively parallel computers [18].

During the post-doctoral position of I.G. Tejada, G. Stoltz, F. Legoll and E. Cancès studied in collaboration with L. Brochard (École des Ponts) the derivation of a concurrent coupling technique to model fractures at the atomistic level by combining a reactive potential with a harmonic approximation; see [50].

6.5. Various topics

A. Bakhta (École des Ponts) and V. Ehrlicher [28] have studied a system of PDEs modeling the cross-diffusion of different atomic species in a crystalline solid thin film during a Physical Vapor Deposition process, coupled with the evolution of the domain as external chemical species fluxes are absorbed at the surface of the solid layer. This model leads to a system of degenerate elliptic cross-diffusion equations. They proved the existence of a global weak solution to this system in arbitrary dimension in the case of a constant domain using analysis tools from gradient flow theory. The existence of a global weak solution in a one-dimensional case with external fluxes was also proved. Under the assumption that this solution is unique, the existence of optimal external fluxes in order to achieve desired concentration profiles of the different species in the thickness of the solid layer at the end of the process was also obtained.

Numerical simulations of crystal defects are necessarily restricted to finite computational domains, supplying artificial boundary conditions that emulate the effect of embedding the defect in an effectively infinite crystalline environment. V. Ehrlicher, in a joint work with C. Ortner (U. of Warwick) and A. Shapeev (Skolkovo Institute of Science and Technology) [39] have studied a mathematical framework within which the accuracy of different types of boundary conditions can be precisely assessed.

T. Lelièvre together with F. Casenave (Safran) and A. Ern (École des Ponts) have proposed in the short note [36] an analysis of the Empirical Interpolation Method which highlights the symmetry played by the two variables (parameter and space variable). A variant of the Empirical Interpolation Method is introduced in order to deal with situations where some observations have to be discarded, and the number of observed values is thus different for the two variables.

In collaboration with P.-L. Lions (Collège de France), C. Le Bris has written an extensive set of lecture notes on parabolic equations with irregular data (initial conditions and parameter coefficients). These lecture notes correspond to joint works between the two authors and to an expanded version of the works by P.-L. Lions specifically exposed in his lectures delivered at Collège de France in 2012–2013. The application of the theory to the specific context of stochastic differential equations with irregular coefficients is also examined.

MATHRISK Project-Team

7. New Results

7.1. Liquidity risk

Participants: Aurélien Alfonsi, Pierre Blanc.

A. Alfonsi and P. Blanc are working on the optimal execution problem when many large traders who modify the market prices. In a previous study, they have developed a price impact model that takes into account an exogenous flow of market orders, in which the optimal execution strategy is known explicitly. This year, they have worked on the practical implementation of this model. Namely, they have proposed an estimation procedure to estimate the model parameters (decay kernel of the price impact and Hawkes kernel for the self excitation of the order flow). They have run this estimation on market data and backtested the optimal execution strategy.

7.2. Backward stochastic (partial) differential equations with jumps, optimal stopping and stochastic control with nonlinear expectation, risk minimization

Participants: Roxana Dumitrescu, Marie-Claire Quenez [(Univ Paris 7)], Arnaud Lionnet, Agnès Sulem.

R. Dumitrescu, M.C. Quenez and A. Sulem have provided a weak dynamic principle for Combined Optimal Stopping/Stochastic Control with \mathcal{E}^f -conditional Expectation. They have investigated the links between generalized Dynkin games and double barriers reflected BSDEs with jumps and also studied mixed generalized Dynkin games in a Markovian framework and associated nonlinear HJB equations with barriers.

In the recent paper [43], they study game options in an imperfect market with default. They extend the results obtained by Kifer [68] in a perfect market model to the case of imperfections taken into account via the nonlinearity of the wealth dynamics. In this framework, the pricing system is expressed as a nonlinear g -expectation/evaluation induced by a nonlinear BSDE with jump. They prove that the superhedging price of a game option coincides with the value function of a corresponding *generalized* Dynkin game expressed in terms of the g -evaluation. They also address the case of ambiguity on the model, - for example an ambiguity on the default probability -, and characterize the superhedging price of a game option as the value function of a *mixed generalized* Dynkin game. They prove the existence of a cancellation time and a trading strategy which allows the seller to be super-hedged, whatever the model is. This study is introduced by the analysis of the simpler case of American options.

In collaboration with Jane Bielagt (Humboldt Univ.) and Gonalo Dos Reis (Univ. of Edimburgh), Arnaud Lionnet investigates in the effects of the social interactions of a finite set of agents on an equilibrium pricing mechanism. They consider an incomplete market where agents invest so as to minimize their risk measure. Here, agents assess risk using convex dynamic risk measures expressed by Backward Stochastic Differential Equations (BSDE). Beside the risk associated with their own economic activity, the agents compare their trading gains to that of the others, and factor this relative performance in the evaluation of their risk/satisfaction. When a derivative product is introduced to complete the market and allow agents to trade a non-financial risk factor (such as temperature), the risk of each agent is lowered, as expected. However, agents then find it in their interest to be more concerned with their relative performance. This leads them to behave more like a herd and this destabilizes the previously stable, purely financial market.

7.3. Systemic risk

Participants: Hamed Amini [EPFL], Andreea Minca [Cornell University], Agnès Sulem, Rui Chen, Romuald Elie.

We study the issue of control of systemic risk in the framework of random graph models. The paper [16] by H. Amini, A. Minca and A. Sulem, provides important insight on the relation between the value of a financial system, connectivity and optimal intervention. More precisely, we consider a core-periphery random financial network in which links lead to the creation of projects in the outside economy but make banks prone to contagion risk. The controller seeks to maximize, under budget constraints, the value of the financial system, defined as the total value of the projects funded. Under partial information on interbank links, revealed in conjunction with the spread of contagion, the optimal control problem is shown to become a Markov decision problem. Our results show that up to a certain connectivity, the value of the financial system increases with connectivity. However, this is no longer the case if connectivity becomes too large. This insight shows that it is far from obvious that connectivity of a core bank should always be brought forward as an argument for priority intervention and it may be sometimes preferable to invest in non-core banks that lend directly to the economy. The natural question remains how to create incentives for the banks to attain an optimal level of connectivity and how to design a guarantee fund that would represent an intervention fund that can be used to maximize the benefits of connectivity. This is under study with the PhD student Rui Chen.

Moreover R. Elie obtained a CVRS PEPS grant on systemic risk modeling with graphs in collaboration with the Inria team COATI and the economic department of Université de Nice.

7.4. Dependence modeling

7.4.1. Estimation of the parameters of a Wishart process

A. Alfonsi with A. Kebaier and C. Rey have studied the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and in some non ergodic cases. In the non ergodic cases, their analysis rely on refined results on the Laplace transform for Wishart processes. This work also extends a recent paper by Ben Alaya and Kebaier on the maximum likelihood estimation for the CIR process.

7.5. Interest rate modeling

A. Alfonsi, E. Palidda and A. Ahdida extend the Linear Gaussian Model (LGM) by replacing the constant covariation matrix by some Wishart dynamics. This extension allows them to generate smile while keeping the affine structure of the model. They have obtained a price expansion around the LGM for Caplet and Swaption prices. They also present a second order discretization scheme that allow them to compute exotic prices within this model.

7.6. Numerical Probability

A. Alfonsi with A. Kohatsu-Higa and M. Hayashi are investigating how to apply the parametric method recently proposed by V. Bally and A. Kohatsu-Higa for reflected SDEs. This method allows them to obtain an unbiased estimator for expectations of general functions of the process.

7.7. Optimal transport

Participant: Benjamin Jourdain.

With J. Corbetta (postdoc financed by the chair financial risks), A. Alfonsi and B. Jourdain study a general formula for the time-derivative of the Wasserstein distance between the time-marginals of two Markov processes. They have checked the validity of this formula for pure-jump Markov processes with a bounded intensity of jumps. They now study the extension to piecewise deterministic Markov processes.

7.8. Multitype sticky particle systems

Participant: Benjamin Jourdain.

B. Jourdain and J. Reygner study multitype sticky particle systems which can be obtained as vanishing noise limits of multitype rank-based diffusions. Rank-based diffusion processes and their multitype generalization permit to reproduce empirical features of stock markets. B. Jourdain and J. Reygner have obtained the optimal rate of convergence as the number of particles grows to infinity of the approximate solution to a diagonal system of hyperbolic conservation laws based on multitype sticky particles.

7.9. Numerical Probability

7.9.1. American option pricing.

Damien Lamberton with M. Pistorius has worked on the approximation of American options by Canadian options, which originated from the work of Peter Carr. This lead them to revise old results on the binomial approximation of the American put. D. Lamberton is also working with M. Zervos on American options involving the maximum of the underlying.

7.9.2. Convergence in total variation of approximation schemes for Markov processes

(V. Bally and PhD student C. Rey [40])

The main issue was to consider very general approximation schemes and to estimate the approximation error for test functions which are just measurable and bounded. It is worth to mention that the input of noise in the approximation schemes is allowed to be quite general, while in the standard approximation schemes for diffusion processes one considers Gaussian input only. In some sense this means that we treat invariance principles as well. We also considered approximation schemes of higher order, as the Victoir Nynomia scheme for example. An important ingredient is an abstract Malliavin calculus for general random variables (which has been settled in previous papers of V. Bally and Lucia Caramellino.

7.9.3. Approximation schemes for Piecewise Deterministic Markov Processes

(V. Bally and PhD student V.Rabiet [39]).

PDMP processes are very popular in many practical fields as biology, chemistry or fiability theory. The main idea is that such a model may present different scales: slow ones and rapid ones. And from a numerical point of view it is extremely difficult to implement algorithms which take care of rapid scales in details. Then the idea is to average the rapid scales (in the spirit of the Central Limit Theorem) and consequently to replace small (and rapid) jumps by a Brownien component. This procedure is already widely used by practitioners. Our work was to derive estimates of the error which is done by this procedure.

7.9.4. Convergence in distribution norms in the Central Limit Theorem

(V. Bally with Lucia Caramellino and Guillaume Poly)

In the classical theory, the convergence which has already been studied is the convergence in total variation (measurable test functions). The main result is the theorem of Prohorov, in the fifties. We have proved that under similar hypothesis (with more finite moments however) one may obtain a much more accurate estimate of the error, in some norms which are close to distribution norms. As a remarkable consequence, we obtained a CLT for the zeros of trigonometric polynomials with random coefficients.

Maxplus Team

7. New Results

7.1. Théorie spectrale max-plus et géométrie métrique/Max-plus spectral theory and metric geometry

7.1.1. Introduction

Participants: Marianne Akian, Stéphane Gaubert, Cormac Walsh.

Étant donné un noyau $a : S \times S \rightarrow \mathbb{R} \cup \{-\infty\}$, on peut lui associer le problème spectral max-plus

$$\sup_{y \in S} a(x, y) + u(y) = \lambda + u(x), \quad \forall x \in S, \quad (1)$$

dans lequel on cherche le vecteur propre $u : S \rightarrow \mathbb{R} \cup \{-\infty\}$ et la valeur propre correspondante $\lambda \in \mathbb{R} \cup \{-\infty\}$. Comme nous l'avons rappelé dans les §3.2 et 3.3, le problème spectral (9) intervient en contrôle ergodique: l'ensemble S est l'espace des états, et l'application $a(x, y)$ fournit le gain associé à la transition $x \rightarrow y$. Le cas où S est fini est classique, l'on a alors un résultat précis de représentation de l'espace propre, à l'aide d'un certain graphe, dit graphe critique. Des résultats existent également lorsque S est compact et que le noyau vérifie certaines propriétés de régularité.

Dans [51], nous avons considéré le cas où S est non compact. Lorsque $\lambda = 0$, l'espace propre est analogue à l'espace des fonctions harmoniques défini en théorie (classique ou probabiliste) du potentiel. En introduisant l'analogue max-plus de la frontière de Martin, nous avons obtenu un analogue de la formule de représentation de Poisson des fonctions harmoniques : toute solution u de (9) peut être représentée sous la forme :

$$u = \sup_{w \in \mathcal{M}_m} w + \mu_u(w), \quad (2)$$

où $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ est l'analogue max-plus de la frontière de Martin minimale (l'ensemble des fonctions harmoniques extrémales normalisées), et où μ_u joue le rôle de la mesure spectrale. Nous avons montré aussi que les éléments de l'espace de Martin minimal peuvent être caractérisés comme les limites de "quasi-géodésiques". La frontière de Martin max-plus généralise dans une certaine mesure la frontière d'un espace métrique construite à partir des horo-fonctions (fonctions de Busemann généralisées), ou horo-frontière. Ces résultats inspirent les travaux des sections suivantes, qui portent sur des cas remarquables d'espaces métriques (§7.1.2) ou sur des applications en théorie des jeux (§7.2.2).

English version

Let the kernel $a : S \times S \rightarrow \mathbb{R} \cup \{-\infty\}$ be given. One may associate the max-plus spectral equation (9), where the eigenvector $u : S \rightarrow \mathbb{R} \cup \{-\infty\}$ and the eigenvalue $\lambda \in \mathbb{R} \cup \{-\infty\}$ are unknown. As we recalled in §3.2 and refmonotone, this spectral problem arises in ergodic optimal control: the set S is the *state space*, and the map $a(x, y)$ is the *transition reward*. The case when S is finite is classical, a precise spectral theorem is known, with a characterisation of the eigenspace in terms of a critical graph. Some results have been shown when S is compact, assuming that the kernel a satisfies some regularity properties.

In [51], we considered the case where S is non-compact. When $\lambda = 0$, the eigenspace is analogous to the set of harmonic functions defined in classical or probabilistic potential theory. By introducing a max-plus analogue of the classical Martin boundary, we obtained an analogue of the Poisson representation of harmonic functions, showing that any solution u of (9) may be represented as in (10) where $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ is a max-plus analogue of the minimal Martin boundary (the set of normalised extremal harmonic functions), and μ_u plays the role of the spectral measure. We also showed that the elements of the minimal Martin boundary can be characterised as limits of certain “almost-geodesics”. The max-plus Martin boundary generalises to some extent the boundary of metric spaces defined in terms of horofunctions (generalised Busemann functions), or horoboundary. These results have inspired the work of the next sections, which deal either with interesting examples of metric spaces (§7.1.2) or with applications to zero-sum games (§7.2.2).

7.1.2. *Isométries de la géométrie de Hilbert/Isometries of the Hilbert geometry*

Participants: Cormac Walsh, Bas Lemmens [Kent University, UK].

Dans nos travaux précédents, nous avons étudié la géométrie de Hilbert (d’un ensemble convexe) en dimension finie, en particulier son horo-frontière et son groupe des isométries. Le chapitre de livre [167] donne une vue d’ensemble de ces travaux. Le cas de la dimension infinie est aussi intéressant, et a été utilisé depuis de nombreuses années en analyse non linéaire. Malgré cela, la géométrie de ces espaces est très peu connue en dimension infinie.

On s’intéresse par exemple au problème suivant. En dimension finie, il est connu que la géométrie de Hilbert est isométrique à un espace normé si et seulement si le convexe est un simplexe. On a montré [38] plus généralement que la géométrie de Hilbert est isométrique à un espace de Banach si et seulement si le convexe est le cône des fonctions positives continues sur un espace topologique compact. Pour cela, on a étudié l’horo-frontière en dimension infinie.

On continue à travailler sur ce sujet avec Bas Lemmens de l’Université de Kent.

English version

Previously, we have been studying the Hilbert geometry in finite dimensions, especially its horofunction boundary and isometry group. The book chapter [167] contains a survey of this work. However, the infinite dimensional case is also interesting, and has been used as a tool for many years in non-linear analysis. Despite this, very little is known about the geometry of these spaces when the dimension is infinite.

An example of a problem in which we are interested is the following. In finite dimension it is known that a Hilbert geometry is isometric to a normed space if and only if it is a simplex. We have shown [38] that, more generally, a Hilbert geometry is isometric to a Banach space if and only if it is the cross-section of a positive cone, that is, the cone of positive continuous functions on some compact topological space. To solve this problem we found it useful to study the horofunction boundary in the infinite-dimensional case.

We are continuing to study similar problems in relation to this topic in collaboration with Bas Lemmens of the University of Kent.

7.1.3. *Croissance des boules dans la géométrie de Hilbert/Volume growth in the Hilbert geometry*

Participants: Cormac Walsh, Constantin Vernicos [Université Montpellier 2].

Avec Constantin Vernicos de l’Université Montpellier 2, nous étudions la croissance du volume de la boule d’une géométrie de Hilbert (d’un ensemble convexe) en fonction du rayon. En particulier, nous étudions l’entropie volumique:

$$\lim_{r \rightarrow \infty} \frac{\log \text{Vol } B(x, r)}{r}, \quad (3)$$

où $B(x, r)$ désigne la boule de centre x et de rayon r , et Vol est une notion de volume particulière, telle que celle définie par Holmes–Thompson ou celle de Busemann. L’entropie ne dépend pas du choix particulier de x , ni de celui du volume. Il est connu que pour l’espace hyperbolique, ou toute géométrie de Hilbert dont la frontière est C^2 et de courbure strictement positive, l’entropie est égale à $n - 1$ lorsque la dimension de l’espace est n , et il a été prouvé récemment que ceci correspond aussi à l’entropie maximale d’une géométrie de Hilbert en dimension n .

Constantin Vernicos a montré que, en dimension 2 et 3, l’entropie volumique d’une géométrie d’Hilbert sur une convexe est égale à l’*approximabilité* de la convexe, ce qui est le taux de croissance exponentielle du nombre de sommets nécessaire pour approximer la convexe par un polytope à ϵ près, quand ϵ diminue.

Ceci motive l’étude de la croissance du volume dans le cas de polytopes. Dans ce cas, la croissance est polynomiale de degré n , plutôt qu’exponentielle, et il est important de comprendre le lien entre le coefficient dominant du polynôme exprimant le volume et la complexité du polytope. Nous avons obtenu une formule pour ce coefficient, laquelle dépend de la structure combinatoire du polytope. Cette formule suggère de définir une nouvelle notion de *approximabilité* en utilisant une quantité combinatoire différente que le nombre de sommets, et d’étudier la relation entre cette *approximabilité* et l’entropie volumique. On pourrait supposer que les deux quantités sont égales, ce qui impliquerait en particulier que l’entropie volumique d’une convexe est égale à celle de son dual.

English version

In a collaboration with Constantin Vernicos of Université Montpellier 2, we are investigating how the volume of a ball in a Hilbert geometry grows as its radius increases. Specifically, we are studying the volume entropy (11) where $B(x, r)$ is the ball with center x and radius r , and Vol denotes some notion of volume, for example, the Holmes–Thompson or Busemann definitions. Note that the entropy does not depend on the particular choice of x , nor on the choice of the volume. It is known that the hyperbolic space, or indeed any Hilbert geometry with a C^2 -smooth boundary of strictly positive curvature, has entropy $n-1$, where n is the dimension, and it has recently been proved that this is the maximal entropy possible for Hilbert geometries of the given dimension.

Constantin Vernicos has shown that, in dimension 2 and 3, the volume entropy of a Hilbert geometry on a convex body is equal to the *approximability* of the body, that is, the exponential rate of growth of the number of vertices needed to approximate the body by a polytope within ϵ , as ϵ decreases.

This motivates studying the volume growth in the polytopal case. Here the growth is polynomial rather than exponential, of degree n , and it is important to know how the constant on front of the highest term depends on the complexity of the polytope. We have a formula for this constant in terms of the combinatorial structure of the polytope. This formula suggests defining a new notion of *approximability* using a different combinatorial quantity from the number of vertices, and studying the relationship between this *approximability* and the volume entropy. One might conjecture that the two quantities are equal, which would imply in particular that the volume entropy of a convex body is equal to that of its dual.

7.1.4. Consensus non-commutatif et contraction d’opérateurs de Kraus/Noncommutative consensus and contraction of Kraus maps

Participants: Stéphane Gaubert, Zheng Qu.

Dans le travail [16], on s’est intéressé à la vitesse de convergence vers l’équilibre d’une itération de la forme $x^{k+1} = T(x^k)$, $x^k \in X$, où T est une application linéaire préservant un cône dans un espace de Banach X , telle que $T(e) = e$, pour un certain vecteur e dans l’intérieur du cône. On s’intéresse aussi à l’itération dans l’espace dual, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, lorsque $\langle y^0, e \rangle = 1$.

Le cas classique est celui où $T(x) = Px$ est un opérateur de Markov. L’itération primale traduit alors la convergence vers le “consensus”, et l’itération duale traduit la convergence de la distribution de probabilité en temps k vers l’état stationnaire. Dans ce cas, le taux de contraction (en un coup) $\kappa(P)$ d’une itération primale, pour la semi-norme de Hilbert $\|z\|_H := \max_i z_i - \min_j z_j$, ainsi que le taux de contraction d’une itération duale, pour la métrique en variation totale, coïncident et sont caractérisés par une formule due à Doeblin et Dobrushin (coefficient d’ergodicité),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^n \min(P_{is}, P_{js}).$$

On a donné ici une généralisation de cette formule au cas d'opérateurs abstraits, qui s'applique en particulier aux opérateurs de Kraus qui interviennent en information quantique. Ces derniers opèrent sur l'espace des matrices symétriques, et sont de la forme

$$T(x) = \sum_k a_k x a_k^* \quad \text{avec} \quad \sum_k a_k a_k^* = I .$$

Dans [114], nous avons étudié des questions de complexité pour les applications de Kraus, montrant en particulier qu'il est NP-dur de vérifier qu'une application de Kraus envoie le cône dans son intérieur.

English version

In [16], we studied the speed of convergence to equilibrium of an iteration of the form $x^{k+1} = T(x^k)$, $x^k \in X$, where T is a linear map preserving a cone in a Banach space X , such that $T(e) = e$, for some vector e in the interior of the cone. We also considered the iteration in the dual space X^* , $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, where $\langle y^0, e \rangle = 1$.

The classical application arises when $T(x) = Px$ is a Markov operator. Then, the primal iteration represents the dynamics of consensus, whereas the dual iteration represents the evolution of the probability distribution as a function of time. Then, the (one-shot) contraction rate $\kappa(P)$ of the primal iteration, with respect to Hilbert's seminorm $\|z\|_H := \max_i z_i - \min_j z_j$, and the contraction rate of the dual iteration, with respect to the total variation metric, coincide, and are characterized by a formula of Doeblin and Dobrushin (ergodicity coefficient),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^n \min(P_{is}, P_{js}).$$

We gave here a generalization of this formula to an abstract operators on a cone. This covers in particular the Kraus maps arising in quantum information theory. The latter maps act on the space of symmetric matrices. They can be written as

$$T(x) = \sum_k a_k x a_k^* \quad \text{with} \quad \sum_k a_k a_k^* = I .$$

In [114], we studied complexity issues related to Kraus maps, and showed in particular that checking whether a Kraus map sends the cone to its interior is NP-hard.

7.2. Algèbre linéaire max-plus, convexité tropicale et jeux à somme nulle/Max-plus linear algebra, tropical convexity and zero-sum games

7.2.1. Polyèdres tropicaux/Tropical polyhedra

Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA], Ricardo Katz [Conicet, Argentine].

On étudie les analogues max-plus ou tropicaux des ensembles convexes. Ceux-ci sont utiles en particulier pour représenter de manière effective les ensembles d'états accessibles de systèmes à événements discrets [9], ils sont aussi apparus récemment en géométrie tropicale, dans toute une série de travaux à la suite de Sturmfels et Develin [97]. Les polyèdres max-plus peuvent aussi être vus comme des limites de déformations de polyèdres classiques, sur lesquels ils donnent un éclairage de nature combinatoire. Toutes ces motivations ont inspiré la recherche d'analogues des résultats fondamentaux d'analyse convexe classique: séparation, projection, points extrémaux, à la suite en particulier de [8].

Dans un travail de X. Allamigeon, S. Gaubert, et E. Goubault [62], [63], on a mis en évidence un critère combinatoire pour la caractérisation des sommets des polyèdres tropicalement convexes. Celui-ci s'exprime à l'aide d'hypergraphes orientés, et de leurs composantes fortement connexes. Ce critère possède la propriété d'être vérifiable en un temps presque linéaire en la taille de l'hypergraphe.

On en déduit un analogue tropical de la méthode de la double description [63] (méthode très utilisée sur les polyèdres classiques, et dûe à Motzkin *et al.* [148]). Cet algorithme permet de calculer les sommets d'un polyèdre défini de façon externe (intersection de demi-espaces ou d'hyperplans tropicaux). Grâce au critère combinatoire précédent, l'algorithme améliore de plusieurs ordres de grandeur les techniques connues jusqu'alors. Ceci est confirmé par de nombreuses expérimentations. Ce travail est motivé par des applications à l'analyse statique [61] et aux systèmes à événements discrets [102], dans lesquelles la manipulation de tels polyèdres est le goulot d'étranglement.

Il est connu qu'un polyèdre tropical peut être représenté comme l'enveloppe convexe d'un ensemble minimal de points et rayons, donnés par ses sommets et ses rayons extrêmes [112]. Dans un travail réalisé par X. Allamigeon et R. Katz [64], et effectué en partie lors de visites de R. Katz à Inria, on étudie la question duale de la caractérisation des représentations minimales par demi-espaces. On montre qu'un polyèdre tropical possède *essentiellement* une unique représentation minimale par demi-espaces, lorsque leurs apex appartiennent au polyèdre. On montre que les apex de ces demi-espaces non-redondants correspondent à certains sommets du complexe tropical introduit par Develin et Sturmfels [97]. On introduit également un critère combinatoire pour l'élimination de demi-espaces redondants à l'aide d'hypergraphes orientés.

Dans un travail de X. Allamigeon et R. Katz [65], nous étudions la tropicalisation des représentations par demi-espaces des polyèdres convexes sur le corps des séries de Puiseux. Nous démontrons ainsi une conjecture de Develin et Yu [98]. Celle-ci assure qu'étant donné un polytope tropical pur, il existe un polytope *relevé* sur les séries de Puiseux, dont les demi-espaces associés aux faces se "tropicalisent" en une représentation par demi-espaces du polytope tropical initial.

Des applications de ces travaux à l'algorithmique, concernant en particulier les jeux répétés, sont discutées dans la Section 7.4.1.

English version

We study the max-plus or tropical analogues of convex sets. These have been used in particular to represent effectively the accessible sets of certain discrete event systems [9]. They also appeared in tropical geometry, following the work of Sturmfels and Develin [97]. Max-plus polyhedra can be thought of as limits of deformations of classical polyhedra, on which they give a combinatorial insight. These motivations have inspired the investigation of analogues of basic results of classical convex analysis: separation, projection, representation by extreme points, following [8].

In a work of X. Allamigeon, S. Gaubert, and E. Goubault [63], we introduce a combinatorial criterion for the characterization of the vertices of tropically convex polyhedra. It is expressed in terms of directed hypergraphs and their strongly connected components. This criterion can be verified in almost linear time in the size of the hypergraph.

This allows to develop a tropical analogue of the double description method [63] (this method is widely used for classical convex polyhedra, and is due to Motzkin *et al.* [148]). This algorithm is able to determine all the vertices of a polyhedron defined externally (intersection of tropical half-spaces of hyperplanes). Thanks to the combinatorial criterion mentioned above, the algorithm improves the existing methods by several orders of magnitude. This is confirmed by several experiments. This is motivated by applications to static analysis [61] and discrete event systems [102], in which computing such polyhedra turns out to be the bottleneck.

It is well-known that a tropical polyhedron can be represented as the convex hull of a minimal set of points and rays, provided by its vertices and extreme rays [112]. In a work of X. Allamigeon and R. Katz [64], partly done during visits of R. Katz at Inria, the dual problem of characterizing the minimal representations by half-spaces is studied. We show that a tropical polyhedron admits *essentially* a unique minimal external representation by half-spaces, provided that their apices belong to the polyhedron. We prove that the apices of

these half-spaces correspond to certain vertices of the tropical complex introduced by Develin and Sturmfels [97]. We also establish a combinatorial criterion allowing to eliminate redundant half-spaces using directed hypergraphs.

In a work of X. Allamigeon and R. Katz [35], we study the tropicalization of the representation by half-spaces of convex polyhedra over the field of Puiseux series. In particular, we prove a conjecture of Develin and Yu [98]. It states that, given a pure tropical polytope, there exists a lifting polytope over Puiseux series, such that the facet-defining half-spaces are “tropicalized” into a representation by half-spaces of the initial polytope.

Some algorithmic applications of this work concerning in particular mean payoff games, will be discussed in Section 7.4.1 .

7.2.2. *Points fixes d’applications monotones homogènes et jeux à somme nulle/Fixed points of order preserving homogeneous maps and zero-sum games*

Participants: Marianne Akian, Stéphane Gaubert, Antoine Hochart.

Pour les jeux répétés à somme nulle, un problème de base est de savoir si le paiement moyen par unité de temps est indépendant de l’état initial. Ici, on définit le paiement moyen directement au moyen de l’opérateur de Shapley (ou de la programmation dynamique) du jeu, lequel préserve l’ordre et commute avec l’addition d’une constante. Dans le cas particulier des jeux à zero joueur, i.e. de chaînes de Markov avec fonctionnelle additive, la solution du problème ci-dessus est fournie par le théorème ergodique. Dans [11], [21], on généralise ce résultat au cas des jeux répétés à espace d’états fini. Cette généralisation est basée sur l’étude de la sous-classe d’opérateurs de Shapley *sans-paiement* (le paiement a lieu seulement le dernier jour), lesquels commutent avec la multiplication par une constante positive. L’intérêt de cette sous-classe est qu’elle inclue la fonction de récession d’un opérateur de Shapley, lorsqu’elle existe. Nous montrons que le paiement moyen est indépendant de l’état initial pour toutes les perturbations des paiements instantanés dépendantes de l’état si, et seulement si, une condition d’ergodicité est vérifiée. Cette dernière est caractérisée par l’unicité, à constante additive près, du point fixe de la fonction de récession de l’opérateur de Shapley, ou, dans le cas particulier des jeux stochastiques à nombre fini d’actions et information parfaite, par une condition d’accessibilité dans un hypergraphe orienté, entre deux sous-ensembles conjugués d’états. On montre aussi que l’ergodicité d’un jeu ne dépend que de la probabilité de transition et qu’elle peut être vérifiée en temps polynomial lorsque le nombre d’états est fixé. Dans [26], on généralise la condition d’accessibilité dans un hypergraphe orienté au cas de jeux avec espaces d’actions arbitraires.

English version

A basic question for zero-sum repeated games consists in determining whether the mean payoff per time unit is independent of the initial state. Here the mean payoff is defined in terms of the Shapley operator (dynamic programming operator) of the game, which is an order preserving map commuting with the addition of a constant. In the special case of “zero-player” games, i.e., of Markov chains equipped with additive functionals, the answer to the above question is provided by the mean ergodic theorem. In [11], [21], we generalize this result to repeated games with a finite state space. This generalization is based on the study of the subclass of *payment-free* Shapley operators (the payment only occurs when the game stops), which are commuting with the multiplication by a positive constant, and which include the recession function of any Shapley operator, when it exists. We show that the mean payoff is independent of the initial state for all state-dependent perturbations of the rewards if and only if an ergodicity condition is satisfied. The latter is characterized by the uniqueness modulo additive constants of the fixed point of the recession function of the Shapley operator, or, in the special case of stochastic games with finite action spaces and perfect information, by a reachability condition involving conjugate subsets of states in directed hypergraphs. We show that the ergodicity condition for games only depends on the support of the transition probability and that it can be checked in polynomial time when the number of states is fixed. In [26], we generalize the above reachability condition to the case of games with arbitrary actions spaces,

7.2.3. *Puissances extérieures tropicales de matrices/Tropical compound matrix identities*

Participants: Marianne Akian, Stéphane Gaubert, Adi Niv.

English version

In [43], [45], we proved some identities on matrices using a weak and strong transfer principles. In the present work, we prove identities on compound matrices in extended tropical semirings. Such identities include analogues to properties of conjugate matrices, powers of matrices and $\text{adj}(A) \det(A)^{-1}$, all of which have implications on the eigenvalues of the corresponding matrices. A tropical Sylvester-Franke identity is provided as well. Even though part of these identities hold over any commutative ring, they cannot be adjusted to semirings with symmetry using the existing weak and strong transfer principles. By reducing these identities to definite matrices, we introduce a transfer principle for formal series, allowing us to infer tropical identities from classical ones. We provide the proofs both via this wider principle and by means of graph theory arguments.

7.2.4. Matrices totalement positives tropicales/Tropical totally positive matrices

Participants: Stéphane Gaubert, Adi Niv.

English version

We investigate totally positive and totally non-negative matrices over the tropical symmetrized semiring. We show these matrices are diagonally dominant, and are determined by their 2×2 minors. We provide the role of the classical double echelon forms in the tropical setting, and find a so called “staircase” form to finite matrices. We establish the connection to the classical sets of totally positive and totally non-negative matrices through the valuation on the field of Puiseux series. In particular, we find the connection to elementary matrix factorization, positivity of eigenvalues and planar networks.

7.2.5. Algèbre supertropicale/Supertropical algebra

Participant: Adi Niv.

English version

Several properties of matrices over the tropical algebra are studied using the supertropical algebra introduced in [126].

The only invertible matrices in tropical algebra are diagonal matrices, permutation matrices and their products. However, the pseudo-inverse A^∇ , defined as $\frac{1}{\det(A)} \text{adj}(A)$, with $\det(A)$ being the tropical permanent, inherits some classical algebraic properties and has some surprising new ones. In [19], defining B and B' to be tropically similar if $B' = A^\nabla B A$, we examine the characteristic (max-)polynomials of tropically similar matrices as well as those of pseudo-inverses. Other miscellaneous results include a new proof of the identity for $\det(AB)$ and a connection to stabilization of the powers of definite matrices.

In a joint work with Louis Rowen (Bar Ilan Univ.) [37], we study the pathology that causes tropical eigenspaces of distinct supertropical eigenvalues of a non-singular matrix A , to be dependent. We show that in lower dimensions the eigenvectors of distinct eigenvalues are independent, as desired. The index set that differentiates between subsequent essential monomials of the characteristic polynomial, yields an eigenvalue λ , and corresponds to the columns of the eigenmatrix $A + \lambda I$ from which the eigenvectors are taken. We ascertain the cause for failure in higher dimensions, and prove that independence of the eigenvectors is recovered in case the “difference criterion” holds, defined in terms of disjoint differences between index sets of subsequent coefficients. We conclude by considering the eigenvectors of the matrix $A^\nabla := \frac{1}{\det(A)} \text{adj}(A)$ and the connection of the independence question to generalized eigenvectors.

In a joint work with Zur Izhakian (University of Aberdeen) and Louis Rowen (Bar Ilan Univ.) [36], extending earlier work on supertropical adjoints and applying symmetrization, we provide a symmetrized supertropical version SL_n of the special linear group, which we partition into submonoids, based on “quasi-identity” matrices, and we display maximal sub-semigroups of SL_n . We also study the monoid generated by SL_n . Several illustrative examples are given of unexpected behavior. We describe the action of elementary matrices on SL_n , which enables one to connect different matrices in SL_n , but in a weaker sense than the classical situation.

7.3. Algèbre max-plus, déformations et asymptotiques /Max-plus algebra, deformations and asymptotic analysis

7.3.1. Introduction

Comme indiqué dans le §3.7 , l'algèbre max-plus est la limite d'une déformation de l'algèbre classique, ou plutôt du semi-corps des réels positifs. Elle peut aussi fournir des estimations de ces déformations, puisque

$$\max(a, b) \leq \epsilon \log(e^{a/\epsilon} + e^{b/\epsilon}) \leq \epsilon \log(2) + \max(a, b) . \quad (4)$$

L'utilisation de ces propriétés a déjà conduit dans le passé aux travaux sur les perturbations de valeurs propres [42], [41], [40], ou sur les grandes déviations [1], [47]. Dans les travaux qui suivent, nous exploitons ces propriétés dans des contextes reliés ou similaires à ceux de nos travaux précédents.

English version

As detailed in §3.7 , max-plus algebra is the limit of a deformation of classical algebra, or more precisely of the semi-field of usual real positive numbers. It can also give estimations for these deformations using for instance (12) . By using these properties, we already obtained some works on singular perturbations of matrix eigenvalues [42], [41], [40], or on large deviations [1], [47]. In the works described below, we are exploiting again these properties in contexts that are related or similar to those of our earlier works.

7.3.2. Méthodes tropicales de localisation de valeurs propres de matrices/Tropical methods for the localisation of matrix eigenvalues

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Dans un travail avec Meisam Sharify [50], on a comparé les modules des valeurs propres d'un polynôme matriciel au moyen des racines tropicales du polynôme obtenu en appliquant une norme donnée aux coefficients. En particulier, on a obtenu des inégalités de type majorisation qui généralisent les bornes obtenues par Polya et Ostrowski dans le cas de polynômes scalaires.

Une partie de la thèse d'Andrea Marchesini, présentée dans [49], montre des inégalités de type majorisation entre les modules des valeurs propres d'une matrice et les valeurs propres tropicales de la matrice de ses modules. En particulier, les majorations généralisent l'inégalité de Friedland [108] concernant le rayon spectral.

Nous avons aussi amélioré et généralisé ces inégalités [27], en appliquant différents changements de variables diagonaux à la matrice complexe initiale, lesquels sont construits à partir des variables duales du problème d'affectation optimale paramétrique construit à partir d'une matrice tropicale associée à la matrice complexe. En particulier, lorsqu'on les applique à une matrice companion par blocs, ces inégalités sont similaires à celles de [50].

English version

In a work with Meisam Sharify [50], we compared the moduli of the eigenvalues of a matrix polynomial to the tropical roots of a polynomial obtained by applying a norm to the coefficients of the original matrix polynomial. In particular, we obtained majorization type inequalities which generalize the bounds of Polya and Ostrowski available for scalar polynomials.

One part of the thesis of Andrea Marchesini, presented in [49], shows majorization type inequalities between the moduli of the eigenvalues of a complex matrix and the tropical eigenvalues of the matrix obtained by applying the modulus entrywise. In particular, the upper bounds generalize the inequality of Friedland [108] concerning the spectral radius. The above inequalities were obtained by using the permanent and tropical analogues of the exterior power of a matrix and by showing (combinatorially) properties of their eigenvalues similar to the ones of usual exterior powers.

We also improved and generalized these inequalities, see [27], by applying to the original complex matrix, different diagonal scalings constructed from the dual variables of the parametric optimal assignment constructed from an associated tropical matrix. In particular, when applied to a block companion matrix, our inequalities are similar to the ones in [50].

7.3.3. Méthodes tropicales pour le calcul numérique de valeurs propres de matrices/Tropical methods for the numerical computation of matrix eigenvalues

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Un des buts de la thèse d'Andrea Marchesini était d'utiliser les résultats de localisation de valeurs propres tels que ceux obtenus ci-dessus pour améliorer la précision des algorithmes de calcul numérique de valeurs propres de matrices ou de polynômes matriciels, en particulier en construisant des changements d'échelle exploitant les calculs tropicaux, à effectuer préalablement à l'appel d'algorithmes classiques comme QZ. Le "changement d'échelle tropical" introduit par Stéphane Gaubert et Meisam Sharify [115] dans le cas de polynôme matriciels quadratiques consiste en un changement de variable multiplicatif de la variable scalaire du polynôme matriciel. Dans la deuxième partie de la thèse d'Andrea Marchesini, en collaboration avec Françoise Tisseur de l'Université de Manchester [22], on considère un changement de variables diagonal du polynôme matriciel construit à partir des variables duales du problème d'affectation optimale paramétrique construit dans l'esprit de [40], [34]. On montre l'intérêt de ces changements d'échelle en terme de conditionnement des valeurs propres, et la supériorité du changement de variables diagonal par rapport au changement d'échelle tropical.

English version

One of goals of the PhD thesis of Andrea Marchesini was to use results on the localisation of eigenvalues like the above ones, to improve the accuracy of the numerical computation of the eigenvalues of a complex matrix or matrix polynomial, in particular by applying scaling methods using tropical techniques, which may be used before calling usual algorithms as QZ. The "tropical scaling" introduced by Stéphane Gaubert and Meisam Sharify [115] in the case of quadratic matrix polynomials consists in a multiplicative scaling of the scalar variable of the matrix polynomial. In the second part of the PhD thesis of Andrea Marchesini, corresponding to a work with Françoise Tisseur from Manchester University [22], we consider a diagonal scaling of the matrix polynomial constructed from the dual variables of the parametric optimal assignment constructed in the same spirit as in [40], [34]. We show the interest of these scaling methods on the eigenvalue conditioning, and the superiority of the diagonal scaling with respect to the tropical scaling.

7.3.4. Tropicalisation du chemin central, et application à la courbure/Tropicalization of the central path and application to the curvature

Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Michael Joswig [TU Berlin].

En optimisation, une classe importante d'algorithmes, dits *de points intérieurs*, consiste à suivre une courbe appelée *chemin central* jusqu'à atteindre la solution optimale. Le chemin central d'un programme linéaire $LP(A, b, c) \equiv \min\{c \cdot x \mid Ax \leq b, x \geq 0\}$ est défini comme l'ensemble des solutions optimales (x^μ, w^μ) des problèmes à barrière logarithmique:

$$\begin{aligned} \text{minimiser} \quad & c \cdot x - \mu \left(\sum_{j=1}^n \log x_j + \sum_{i=1}^m \log w_i \right) \\ \text{sous les contraintes} \quad & Ax + w = b, \quad x > 0, \quad w > 0 \end{aligned}$$

Les performances d'un algorithme de point intérieur sont intimement liées à la forme du chemin central. En particulier, la courbure mesure de combien un chemin diffère d'une ligne droite. Intuitivement, un chemin central à forte courbure devrait être plus difficile à approximer par des segments de droites, ce qui suggère davantage d'itérations des algorithmes de points intérieurs. La courbure totale du chemin central a été étudiée par Dedieu, Malajovich et Shub [94] à travers le théorème de Bezout dans le cas multihomogène, et par De

Loera, Sturmfels and Vinzant [93] à l'aide de la théorie des matroïdes. Ces deux travaux fournissent une borne supérieure en $O(n)$ sur la courbure totale moyenne sur l'ensemble des régions formées par l'arrangement d'hyperplans en dimension n . Le cube de Klee-Minty redondant de [100] et le "serpent" de [99] sont des instances qui montrent que la courbure totale peut être de l'ordre de $\Omega(m)$ pour un polytope défini par m inégalités.

Dans un travail de X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, nous avons étudié la tropicalisation du chemin central. Le *chemin central tropical* est défini comme la limite logarithmique des chemins centraux d'une famille paramétrique de programmes linéaires $\text{LP}(A(t), b(t), c(t))$, où les entrées $A_{ij}(t)$, $b_i(t)$ and $c_j(t)$ sont des fonctions définissables dans une structure o-minimale appelée *corps de Hardy*.

Une première contribution a été de fournir une caractérisation entièrement géométrique du chemin central tropical. Nous avons montré que le centre analytique est donné par le plus grand élément de l'ensemble des points tropicaux admissibles. De plus, tout point du chemin central tropical coïncide avec le plus grand élément de l'ensemble admissible tropical intersecté avec un ensemble de sous-niveau de la fonction de coût tropicale.

Grâce à cette caractérisation, nous avons réfuté l'analogue continu de la conjecture de Hirsch proposé par Deza, Terlaky et Zinchenko. Ainsi, nous avons construit une famille de programmes linéaires définis par $3r + 4$ inégalités in dimension $2r + 2$, où le chemin central a une courbure totale en $\Omega(2^r)$. Cette famille est obtenue en relevant des programmes linéaires tropicaux qui proviennent d'une construction de Bezem, Nieuwenhuis et Rodríguez-Carbonell [73]. Afin d'estimer la courbure dans ce contre-exemple, nous introduisons une notion d'angle combinatoire, qui est de nature tropicale. Cela nous permet de définir un analogue combinatoire de la courbure totale qui fournit un minorant de la courbure totale classique.

Ces résultats sont rassemblés dans le document [60].

English version

In optimization, path-following interior point methods are driven to an optimal solution along a trajectory called the central path. The *central path* of a linear program $\text{LP}(A, b, c) \equiv \min\{c \cdot x \mid Ax \leq b, x \geq 0\}$ is defined as the set of the optimal solutions (x^μ, w^μ) of the barrier problems:

$$\begin{aligned} \text{minimize} \quad & c \cdot x - \mu \left(\sum_{j=1}^n \log x_j + \sum_{i=1}^m \log w_i \right) \\ \text{subject to} \quad & Ax + w = b, \quad x > 0, \quad w > 0 \end{aligned}$$

The performance of an interior point method is tightly linked to the shape of its central path. In particular, the curvature measures how far a path differs from a straight line. Intuitively, a central path with high curvature should be harder to approximate with line segments, and thus this suggests more iterations of the interior point methods. The total curvature of the central path has been studied by Dedieu, Malajovich and Shub [94] via the multihomogeneous Bézout Theorem and by De Loera, Sturmfels and Vinzant [93] using matroid theory. These two papers provide an upper bound of $O(n)$ on the total curvature averaged over all regions of an arrangement of hyperplanes in dimension n . The redundant Klee-Minty cube of [100] and the "snake" in [99] are instances which show that the total curvature can be in $\Omega(m)$ for a polytope described by m inequalities. By analogy with the classical Hirsch conjecture, Deza, Terlaky and Zinchenko [99] conjectured that $O(m)$ is also an upper bound for the total curvature.

In a work of X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, we have studied the tropicalization of the central path. The *tropical central path* is defined as the logarithmic limit of the central paths of a parametric family of linear programs $\text{LP}(A(t), b(t), c(t))$, where the entries $A_{ij}(t)$, $b_i(t)$ and $c_j(t)$ are definable functions in an o-minimal structure called the *Hardy field*.

A first contribution is to provide a purely geometric characterization of the tropical central path. We have shown that the tropical analytic center is the greatest element of the tropical feasible set. Moreover, any point of the tropical central path is the greatest element of the tropical feasible set intersected with a sublevel set of the tropical objective function.

Thanks to this characterization, we disprove the continuous analog of the Hirsch conjecture proposed by Deza, Terlaky and Zinchenko, by constructing a family of linear programs with $3r + 4$ inequalities in dimension $2r + 2$ where the central path has a total curvature in $\Omega(2^r)$. This family is gotten by lifting tropical linear programs which come from a construction of Bezem, Nieuwenhuis and Rodríguez-Carbonell [73]. In order to estimate the curvature in this counter example, we introduce a notion of a combinatorial angle, which is tropical in nature. This allows us to define a combinatorial analogue of the total curvature which provides a lower bound for the classical total curvature.

These results are gathered in the preprint [60].

7.3.5. *Etude des ensembles semi-algébriques tropicaux/Tropicalization of semi-algebraic sets*

Participants: Xavier Allamigeon, Stéphane Gaubert, Mateusz Skomra.

Suite à son stage de M2, Mateusz Skomra a débuté en octobre 2015 une thèse sous la direction de Xavier Allamigeon et Stéphane Gaubert portant sur la tropicalisation des ensembles semi-algébriques. Cette thèse est financée par une bourse de la région Ile-de-France. La thèse de Mateusz vise en particulier à étudier l'analogie tropical de la programmation sur le cône des matrices positives, dans le but d'obtenir de nouveaux algorithmes ou résultats de complexité pour des sous-classes de problèmes.

English version

Following his M2 internship, Mateusz Skomra has started in October 2015 a PhD thesis under the supervision of Xavier Allamigeon and Stéphane Gaubert, on the tropicalization of semi-algebraic sets. This thesis is funded by a grant from Ile-de-France. One goal is to study the tropical analogue of semidefinite programming, in order to define new algorithms or to establish new complexity results for some subclasses of problems.

7.4. Algorithmes/Algorithms

7.4.1. *Algorithmique des polyèdres tropicaux/Algorithmics of tropical polyhedra*

Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Michael Joswig [TU Berlin].

Dans un travail de X. Allamigeon, P. Benchimol, S. Gaubert et M. Joswig [13], nous avons défini un analogue tropical de l'algorithme du simplexe qui permet de résoudre les problèmes de *programmation linéaire tropicale*, i.e.

$$\begin{aligned} & \text{minimiser} && \max_{1 \leq j \leq n} c_j + x_j \\ & \text{sous les contraintes} && \max(\max_{1 \leq j \leq n} (a_{ij}^+ + x_j), b_i^+) \geq \max(\max_{1 \leq j \leq n} (a_{ij}^- + x_j), b_i^-), \quad i = 1, \dots, m \\ & && x \in (\mathbb{R} \cup \{-\infty\})^n \end{aligned} \quad (5)$$

où les entrées du programme a_{ij}^\pm, b_i^\pm, c_j sont à valeur dans $\mathbb{R} \cup \{-\infty\}$. Ces problèmes sont intimement liés à la résolution de jeux répétés à somme nulle, puisque résoudre un jeu à paiement moyen déterministe est équivalent à déterminer si un problème de programmation linéaire admet un point réalisable [44].

Comme son homologue usuel, le simplexe tropical pivote entre des points de base (tropicaux), jusqu'à atteindre l'optimum du programme linéaire. La différence fondamentale avec l'algorithme du simplexe classique est que le pivotage est réalisé de manière purement combinatoire, en s'appuyant sur des descriptions locales du polyèdre tropical défini par les contraintes à l'aide d'(hyper)graphes orientés. Ceci nous a permis de prouver que *l'étape de pivotage (incluant le calcul des coûts réduits) a la même complexité en temps que dans l'algorithme classique, i.e. $O(n(m+n))$* . Ceci est d'autant plus inattendu que la structure des arêtes tropicales entre deux points de base sont géométriquement plus complexes (elles sont constituées de plusieurs segments de droite, jusqu'à n).

Le simplexe tropical a la propriété d'être fortement corrélé avec l'algorithme du simplexe classique. Grâce au principe de Tarski, le simplexe usuel peut être transposé tel quel sur des programmes linéaires dont les coefficients en entrée sont non plus des réels, mais sur le corps $\mathbb{R}\{\{t\}\}$ des séries de Puiseux généralisées en une certaine indéterminée t , *i.e.* des objets de la forme :

$$c_{\alpha_1} t^{\alpha_1} + c_{\alpha_2} t^{\alpha_2} + \dots \quad (6)$$

où les α_i sont des réels, les coefficients c_{α_i} sont des réels non-nuls, et où la séquence des $\alpha_1, \alpha_2, \dots$ est strictement croissante et soit finie, soit non-bornée. L'opposé du plus petit exposant de la série, $-\alpha_1$, est appelé *valuation* de la série. Un programme linéaire tropical est dit *relevé* en un problème linéaire sur $\mathbb{R}\{\{t\}\}$, si la valuation des coefficients en entrée de ce dernier sont égaux aux coefficients du problème tropical. Dans nos travaux, nous avons établi la correspondance suivante entre le simplexe usuel et le simplexe tropical : *pour tout programme linéaire tropical générique, l'algorithme du simplexe tropical trace l'image par la valuation du chemin sur l'algorithme du simplexe usuel sur n'importe quel relèvement du programme tropical dans $\mathbb{R}\{\{t\}\}$.*

Les résultats présentés ci-dessus sont rassemblés dans l'article [13]. Ils ont fait l'objet de plusieurs présentations en conférence [54], [55] [59].

Ces résultats ouvrent la possibilité de relier la complexité du l'algorithme du simplexe usuel avec celles des jeux déterministes. Pour ces derniers, on sait seulement que leur résolution est dans la classe de complexité $NP \cap coNP$, et on ignore s'il existe un algorithme de complexité polynomiale. De façon similaire, on ne sait pas caractériser de façon précise la complexité de l'algorithme du simplexe usuel. Celle-ci dépend fortement de la règle de pivotage utilisée, et il existe des problèmes sur lesquelles de nombreuses règles de pivotage ont une complexité exponentielle. L'existence d'une règle de pivotage qui permettrait au simplexe de terminer en temps polynomial sur n'importe quelle instance est encore aujourd'hui une question ouverte.

Dans un deuxième travail, nous avons relié les deux problèmes ouverts précédents, grâce à l'algorithme du simplexe tropical. Nous avons en effet exhibé une classe de règles de pivotage, dites *combinatoires*, et avons montré qu'elles satisfont la propriété suivante : *s'il existe une règle de pivotage combinatoire qui permet de résoudre tout problème de programmation linéaire usuel en temps polynomial, alors on peut résoudre les jeux à paiement moyen en temps (fortement) polynomial.* Le terme *combinatoire* fait référence au fait que la règle est définie en fonction du signe des mineurs de la matrice des coefficients du problème linéaire. Ce résultat est décrit dans l'article [56], et a été présenté dans plusieurs conférences [57], [58].

Enfin, dans un travail de X. Allamigeon, P. Benchimol et S. Gaubert [53], nous avons étendu les résultats aux règles de pivotage *semi-algébriques*, classe incluant la règle dite du *shadow-vertex*. Celle-ci est connue pour avoir fourni plusieurs bornes de complexité moyenne et lisse sur l'algorithme du simplexe. Nous avons donc tropicalisé l'algorithme du simplexe shadow-vertex, et nous avons montré que cet algorithme permet de résoudre les jeux à paiement moyen en temps polynomial en moyenne.

English version

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and M. Joswig, we introduced a tropical analogue of the simplex algorithm, allowing one to solve problems of *tropical linear programming*, which are of the form (13), where the coefficients of the program, a_{ij}^\pm, b_i^\pm, c_j take their values in the max-plus semiring $\mathbb{R} \cup \{-\infty\}$. These problems are closely related to mean payoff games, as solving a game of this kind is equivalent to determine whether a tropical linear program admits a feasible point [44].

Like the classical simplex algorithm, the tropical simplex algorithm performs pivoting operations between basis points, until it reaches the optimum. The main discrepancy with the classical algorithm is that the pivoting is now a purely combinatorial operation, which is performed by using a local description of the polyhedron by a directed hypergraph. This allowed us to show that *a tropical pivoting step (including computing reduced costs) has the same complexity as in the classical simplex algorithm, i.e. $O(n(m+n))$.* This is all the more

surprising as the tropical edge between two given points has a geometrically more complex structure in the tropical case (it is constituted of up to n ordinary line segments).

The tropical simplex algorithm turns out to be closely related to the classical one. Thanks to Tarski's principle, the latter is also valid for linear programs over the field $\mathbb{R}\{\{t\}\}$ of generalized Puiseux series in an indeterminate t . These series are of the form (14), where the α_i are real numbers, the coefficients c_{α_i} are non-zero reals, and the sequence $\alpha_1, \alpha_2, \dots$ is strictly increasing and either finite or unbounded. The opposite of the smallest exponent of the series, $-\alpha_1$, is called *valuation*. A tropical linear program is said to be *lifted* to a linear program over $\mathbb{R}\{\{t\}\}$ if the valuation of the coefficients of the latter are sent to the coefficients of the former by the valuation. We showed the following relation between the classical simplex algorithm and its tropical analogue: *for all generic tropical linear program, the tropical simplex algorithm computes the image by the valuation of the path of the classical simplex algorithm, applied to any lift in $\mathbb{R}\{\{t\}\}$ of the original program.*

These results are gathered in the article [13]. They have been presented in several conferences [54], [55] [59].

They allow one to relate the complexity of the classical simplex algorithm with the complexity of mean payoff games. The latter is unsettled, these games are known to be in the class $\text{NP} \cap \text{coNP}$ but it is not known whether they can be solved in polynomial time. Basic complexity issues regarding the classical simplex algorithm are also unsettled: its execution time depends on the pivoting rule, and many pivoting rules have been shown to have exponential worst case behaviors. The existence of a pivoting rule leading the simplex to terminate in polynomial time is still an open question. In a second work, we related these two open questions, via the tropical simplex algorithm. We identified a class of pivoting rules, which are said to be *combinatorial*, and show that they have the following property: *if there is a combinatorial pivoting rule allowing one to solve every classical linear programming problem in polynomial time, then, mean payoff games can be solved in (strongly) polynomial time.* By *combinatorial*, we mean that the rule depends only of the coefficients of the system through the signs of minors of the coefficients matrix. This result is given in the article [56]. It has been presented to the conferences [57], [58].

Finally, in a work of X. Allamigeon, P. Benchimol and S. Gaubert [53], we extended the latter results to *semi-algebraic* pivoting rules, which include the so-called *shadow-vertex* rule. This rule has been exploited in the literature to establish several average-case and smooth complexity bounds on the simplex algorithm. We tropicalized the shadow-vertex simplex algorithm, and showed that it solves mean payoff games in polynomial time on average.

7.4.2. *Approximation max-plus de fonctions valeurs et équations de Riccati généralisées/Max-plus approximation of value functions and generalized Riccati equations*

Participants: Stéphane Gaubert, Zheng Qu, Srinivas Sridharan.

Le travail de thèse de Zheng Qu, supervisée par S. Gaubert et S. Tang, a porté sur le développement de méthodes tropicales en programmation dynamique approchée [154]. Celle-ci permettent d'atténuer la malédiction de la dimension, pour certaines classes de problèmes de contrôle optimal.

Un développement de ce travail est paru dans [17], où il est montré qu'une classe de relaxations convexes introduites par Sridharan et al. pour traiter numériquement un problème de contrôle quantique sont en fait exactes (pas de saut de relaxation).

English version

The PhD work of Zheng Qu, supervised by S. Gaubert and S. Tang, dealt with the développement of tropical methods in approximate dynamic programming [154]. These allow one to attenuate the curse of dimensionality for certain optimal control problems.

A development of this work appeared in [17]. It is shown there that a class of convex relaxations introduced Sridharan et al. to solve numerically some quantum control problem is exact.

7.4.3. Approximation probabiliste d'équations d'Hamilton-Jacobi-Bellman et itération sur les politiques

Participants: Marianne Akian, Eric Fodjo.

La thèse d'Eric Fodjo traite de problèmes de contrôle stochastique (de diffusions) issus en particulier de problèmes de gestion de portefeuille avec coûts de transaction. La programmation dynamique conduit à une équation aux dérivées partielles d'Hamilton-Jacobi-Bellman, sur un espace de dimension au moins égale au nombre d'actifs risqués. La malédiction de la dimension ne permet pas de traiter numériquement ces équations en dimension grande (supérieure à 5). On se propose d'aborder ces problèmes avec des méthodes numériques associant itération sur les politiques, discrétisations probabilistes, et discrétisations max-plus, afin d'essayer de monter plus en dimension. Une autre piste est de remplacer l'itération sur les politiques par une approximation par des problèmes avec commutations optimales.

Nous considérons actuellement des équations d'Hamilton-Jacobi-Bellman fortement non-linéaires associées à des problèmes de contrôle de diffusions faisant intervenir un contrôle discret (prenant un nombre fini de valeurs) et éventuellement un contrôle continu. On construit un algorithme numérique probabiliste de faible complexité, en combinant les propriétés de distributivité idempotente obtenues par McEneaney, Kaise et Han [128], [145] pour le même type d'équations et la méthode numérique probabiliste proposée par Fahim, Touzi et Warin [104] pour résoudre des équations d'Hamilton-Jacobi-Bellman fortement non-linéaires, lorsque la volatilité ne varie pas trop.

English version

The PhD thesis of Eric Fodjo concerns stochastic control problems obtained in particular in the modelisation of portfolio selection with transaction costs. The dynamic programming method leads to a Hamilton-Jacobi-Bellman partial differential equation, on a space with a dimension at least equal to the number of risky assets. Curse of dimensionality does not allow one to solve numerically these equations for a large dimension (greater to 5). We propose to tackle these problems with numerical methods combining policy iterations, probabilistic discretisations, max-plus discretisations, in order to increase the possible dimension. Another solution is to replace policy iterations by an approximation with optimal switching problems.

Our current work concerns fully nonlinear Hamilton-Jacobi-Bellman equations associated to diffusion control problems with finite horizon involving a finite set-valued (or switching) control and possibly a continuum-valued control. We construct a lower complexity probabilistic numerical algorithm by combining the idempotent expansion properties obtained by McEneaney, Kaise and Han [128], [145] for solving such problems with a numerical probabilistic method such as the one proposed by Fahim, Touzi and Warin [104] for solving some fully nonlinear parabolic partial differential equations, when the volatility does not oscillate too much.

7.5. Applications

7.5.1. Introduction

Nous présentons maintenant plusieurs travaux de nature appliquée, touchant à des domaines variés, dans lesquels nous exploitons certaines des techniques mathématiques présentées précédemment, et particulièrement celles qui relèvent de la théorie de Perron-Frobenius non-linéaire et de la convexité tropicale. Ces applications utilisent aussi des techniques d'algèbre linéaire ou d'optimisation convexe.

English version

In this section, we describe several applied works in which we use some of the theoretical tools developed by the team, including non-linear Perron-Frobenius theory and tropical convexity. Some of these applications also make an intensive use of linear algebraic and convex programming methods.

7.5.2. Preuve formelle d'inégalités non-linéaires/Formal proofs of non-linear inequalities

Participants: Xavier Allamigeon, Stéphane Gaubert, Victor Magron, Benjamin Werner [LIX].

La thèse de Victor Magron [140], dirigée par Benjamin Werner, codirigée par Stéphane Gaubert et Xavier Allamigeon, a porté sur la certification de bornes inférieures de fonctions multivariées à valeurs réelles, définies par des expressions semi-algébriques ou transcendentes, et sur la preuve de validité de celles-ci au moyen de certificats dans l'assistant de preuves Coq.

L'un des développements de ce travail est paru dans [18],

English version

The PhD work of Victor Magron [140], supervised by Benjamin Werner, and cosupervised by Stéphane Gaubert and Xavier Allamigeon, dealt with the certification of lower bounds for multivariate functions, defined by semi-algebraic or transcendental expressions, and their correctness proof through certificates checked in the Coq proof assistant. A development of this work appeared in [18]

7.5.3. Géométrie de l'ordre de Loewner et application au calcul d'invariants quadratiques en analyse statique de programme/Geometry of the Loewner order and application to the synthesis of quadratic invariants in static analysis of program

Participants: Xavier Allamigeon, Stéphane Gaubert, Éric Goubault [LIX], Sylvie Putot [LIX], Nikolas Stott.

Nous introduisons un nouveau domaine abstrait numérique reposant sur les ellipsoïdes pour la vérification formelle de systèmes linéaires switchés. La nouveauté de ce domaine ne réside pas dans l'utilisation des ellipsoïdes en tant qu'abstraction, mais dans le fait que nous dépassons deux difficultés clés qui ont jusqu'à maintenant limité l'utilisation des ellipsoïdes en interprétation abstraite. La première difficulté est que l'ensemble des ellipsoïdes ne constitue pas un treillis. Par conséquent, il n'y a pas a priori de choix canonique pour l'abstraction de l'union de deux ensembles, ce qui rend l'analyse moins prévisible (elle dépend du choix de "bonnes" bornes supérieures). La seconde difficulté est que les travaux récents utilisant les ellipsoïdes reposent sur des méthodes à base d'inégalités linéaires matricielles (LMI). Ces dernières sont efficaces sur des exemples de taille modérée, mais elles sont limitées par la complexité des algorithmes de points intérieurs. Ceux-ci ne passent pas aussi bien à l'échelle dans le cas des LMI que dans le cas de la programmation linéaire ou du second ordre.

Plus précisément, nous réduisons la question de l'abstraction de l'union de deux ensembles par une ellipsoïde à la sélection d'une borne inférieure de deux matrices positives pour l'ordre de Löwner. Nous montrons qu'il existe une unique procédure de sélection ayant la propriété d'être invariante par transformation linéaire des variables de programmes. Nous montrons que la borne inférieure ainsi sélectionnée peut être calculée en $O(n^3)$ opérations arithmétiques. Nous établissons aussi que cette borne inférieure coïncide avec l'ellipsoïde de volume minimal, si bien que, de façon surprenante, deux approches distinctes et naturelles mènent à la même sélection. Par ailleurs, nous montrons qu'un invariant ellipsoïdal peut être calculé de manière efficace. Notre algorithme est fondé sur une généralisation non-linéaire de l'algorithme power, utilisé habituellement pour calculer la plus grande valeur propre d'une matrice. Nous illustrons notre approche en l'appliquant à des exemples de systèmes linéaires switchés. Nous montrons que l'algorithme power conduit à des gains en temps de calcul très importants par rapport aux méthodes de type LMI, au prix d'une perte de précision limitée.

Ce travail est décrit dans l'article [29], qui a reçu le prix du meilleur article à la conférence EMSOFT 2015.

English version

We introduce a new numerical abstract domain based on ellipsoids designed for the formal verification of switched linear systems. The novelty of this domain does not consist in the use of ellipsoids as abstractions, but rather in the fact that we overcome two key difficulties which so far have limited the use of ellipsoids in abstract interpretation. The first issue is that the ordered set of ellipsoids does not constitute a lattice. This implies that there is a priori no canonical choice of the abstraction of the union of two sets, making the analysis less predictable as it relies on the selection of good upper bounds. The second issue is that most recent works using on ellipsoids rely on LMI methods. The latter are efficient on moderate size examples but they are inherently limited by the complexity of interior point algorithms, which, in the case of matrix inequality problems, do not scale as well as for linear programming or second order cone programming problems.

In more details, we reduce the question of abstracting by an ellipsoid the union of two sets to the selection of a minimal upper bound of two positive semidefinite matrices with respect to the Löwner order. We show that there is a unique selection procedure which has the property of being invariant with respect to linear transformations of the program variables. We show that the minimal upper bound can be computed with the same cost as performing a Cholesky decomposition, i.e., in $O(n^3)$ arithmetic operations. We also establish that it coincides with the minimal volume ellipsoid, so that, surprisingly, two distinct natural approaches lead to the same choice of selection. Moreover, we show that an invariant ellipsoid can be computed by a scalable algorithm. This is based on a non-linear generalization of the power algorithm which is classically used to compute the dominant eigenvalue of a matrix. We illustrate our approach by applying it to examples of switched linear systems. We show that the power iteration leads to important speedups by comparison with LMI based methods, at the price of a limited loss of precision.

This work is described in the article [29], and won the best paper award at the conference EMSOFT 2015.

7.5.4. *Optimisation de l'affectation temps réel des moyens de secours des pompiers/Optimization of the real time assignment of firemen vehicles*

Participants: Marianne Akian, Xavier Allamigeon, Vianney Boeuf, Stéphane Gaubert, Stéphane Raclot [BSPP].

La thèse de Vianney Boeuf est effectuée en partenariat avec la Brigade de sapeurs pompiers de Paris (BSPP). Elle est motivée par l'optimisation des moyens de secours, en incluant les questions de dimensionnement et d'affectation temps réel des moyens. On s'intéresse en particulier à l'affectation des engins et véhicules de secours, éventuellement empruntés à différentes casernes. Ce travail intervient en complément du travail de l'équipe au sein du projet ANR Democrite, qui porte sur l'évaluation du risque en milieu urbain.

English version

The PhD work of Vianney is carried out with the Brigade of Paris Firemen (BSPP). It is motivated by the issue of optimization of emergency resources, including the real time dynamic assignment of engines or emergency vehicles. This work is carried out in complement to the ANR project Democrite, dealing with risk evaluation in urban environment.

7.5.5. *Analyse de performance d'un centre de réception des appels d'urgence fondée sur les systèmes polynomiaux tropicaux/Performance evaluation of an emergency call center based on tropical polynomial systems*

Participants: Xavier Allamigeon, Vianney Boeuf, Stéphane Gaubert, Stéphane Raclot [BSPP], Régis Rebol [PP].

Ce travail a pris sa source dans un problème d'évaluation de performance et de dimensionnement, posé par Régis Rebol (Préfecture de Police), portant sur l'analyse de l'évolution projetée de la procédure de traitement des appels d'urgence (17-18-112). Ce travail a aussi bénéficié de l'appui du LtL Stéphane Raclot (BSPP).

Il a amené à la contribution suivante. Nous introduisons une méthode algébrique qui permet d'analyser les performances de systèmes mettant en jeu des priorité et modélisés par des réseaux de Petri. Cette méthode s'applique à la classe de réseaux de Petri dans lesquels les places peuvent être partitionnées en deux catégories : le routage dans certaines places est sujet à des règles de priorité, tandis que le routage dans les autres places est à choix libre.

Nous montrons que les variables compteurs, qui déterminent le nombre de tirage des transitions comme une fonction du temps, sont les solutions d'un système dynamique linéaire par morceaux. Par ailleurs, nous établissons que dans le modèle fluide, les régimes stationnaires sont précisément les solutions d'un ensemble d'équations linéaires par morceaux et lexicographiques, qui constituent un système polynomial sur un semi-corps tropical (min-plus) de germes.

En substance, ce résultat montre que calculer les régimes stationnaires se réduit à la résolution d'un système polynomial tropical. Ceci est l'un des problèmes de base en géométrie tropicale. Cette interprétation fournit des informations sur la nature des solutions ainsi que des algorithmes. En particulier, l'approche tropicale permet de déterminer les différentes phases de congestion du système.

Nous appliquons ensuite cette approche à un cas d'étude relié au projet actuel de la Préfecture de Police de Paris et la Brigade de sapeurs-pompiers de Paris sur la mise en place d'une nouvelle organisation de réception des appels 17/18/112 sur Paris et sa petite couronne. Nous introduisons pour cela un modèle simplifié d'organisation de réception des appels, et nous nous concentrons sur l'analyse d'une fonctionnalité clé de cette organisation : la procédure de réception des appels à deux niveaux. Les opérateurs de niveau 1 reçoivent les appels, les qualifient en fonction de leur urgence, gèrent les appels non-urgents, et transfèrent les appels urgents à des opérateurs de niveau 2 spécialisés et qui complètent l'instruction. Nous résolvons le système d'équations polynomiales tropicales correspondant, et obtenons un calcul explicite des différents phases de congestion en fonction du rapport entre les nombres d'opérateurs de niveau 2 et 1. Nos résultats analytiques sont obtenus pour le modèle fluide. Cependant, ils sont confirmés par des simulations dans lesquelles la sémantique initiale du réseau de Petri est utilisée.

Ce travail a fait l'objet de la publication [28] à la conférence FORMATS 2015.

English version

This work arose from a question raised by Régis Rebol (Préfecture de Police), regarding the analysis of the projected evolution of the treatment of emergency calls (17-18-112). This work also benefited from the help of LtL Stéphane Raclot (BSPP).

It led to the following contribution. We introduce an algebraic approach which allows to analyze the performance of systems involving priorities and modeled by timed Petri nets. Our results apply to the class of Petri nets in which the places can be partitioned in two categories: the routing in certain places is subject to priority rules, whereas the routing at the other places is free choice.

We show that the counter variables, which determine the number of firings of the different transitions as a function of time, are the solutions of a piecewise linear dynamical system. Moreover, we establish that in the fluid model, the stationary regimes are precisely the solutions of a set of lexicographic piecewise linear equations, which constitutes a polynomial system over a tropical (min-plus) semifield of germs.

In essence, this result shows that computing stationary regimes reduces to solving tropical polynomial systems. Solving tropical polynomial systems is one of the most basic problems of tropical geometry. The latter provides insights on the nature of solutions, as well as algorithmic tools. In particular, the tropical approach allows one to determine the different congestion phases of the system.

We apply this approach to a case study relative to the current project led by Préfecture de Police de Paris (PP), involving the Brigade de sapeurs-pompiers de Paris (BSPP), of a new organization to handle emergency calls to Police (number 17), Firemen (number 18), and untyped emergency calls (number 112), in the Paris area. To this purpose, we introduce a simplified model of emergency call center, and we concentrate on the analysis of an essential feature of the organization: the two level emergency procedure. Operators at level 1 initially receive the calls, qualify their urgency, handle the non urgent ones, and transfer the urgent cases to specialized level 2 operators who complete the instruction. We solve the associated system of tropical polynomial equations and arrive at an explicit computation of the different congestion phases, depending on the ratio of the numbers of operators of level 2 and 1. Our analytical results are obtained only for the approximate fluid model. However, they are confirmed by simulations in which the original semantics of the Petri nets (with integer firings) is respected.

This work has been published in the proceedings of the conference FORMATS 2015 [28].

7.5.6. Tarification du tarif des données dans les réseaux mobiles/Smart Data Pricing

Participants: Marianne Akian, Mustapha Bouhtou [Orange Labs], Jean-Bernard Eytard.

Le travail de PhD de Jean-Bernard Eytard, qui a démarré en Octobre, concerne l'optimisation de la tarification des données dans les réseaux mobiles.

English version

The PhD work of Jean-Bernard Eytard, which started in October, concerns the optimal pricing of data traffic in mobile networks.

MCTAO Project-Team

6. New Results

6.1. Optimal control for quantum systems and applications to MRI

Participants: Bernard Bonnard, Thierry Combot [Université de Bourgogne, IMB], Alain Jacquemard [Université de Bourgogne, IMB], Dominique Sugny [Université de Bourgogne, LIC].

Important results have been obtained in this area that we detail next :

- A complete solution to the time minimal control of a chain of three spins with Ising coupling which is a toy example applicable to quantum computing [42], [3].
- Optimal control of an ensemble of spins systems with application to MRI : this work is performed in the framework of the ANR project Explosys , based on our previous results in the contrast problem in Nuclear Magnetic Imaging. In relation with the laboratory Creatis (Insa Lyon) and TUM (S. Glaser) the objective is to design robust pulses control, with respect to the relaxation parameters and the B0 and B1 inhomogeneities. The computations are intricate from both the numerical point of view and exact computations. From this second point a systematic study of the controlled Bloch equation has been initiated using exact computer algebraic method in relation with the Inria project-team POLSYS .

6.2. Controllability and Optimal control at Low Reynolds number

Participants: Piernicola Bettiol [Université de Bretagne Occidentale (Brest)], Bernard Bonnard, Laetitia Giraldi, Pierre Martinon [project-team COMMANDS], Jean-Baptiste Pomet, Jérémy Rouot.

This new area is somehow connected to the recent recruitment of L. Giraldi (CR2) in the Mc Tao team. The problem under study is to design strokes for swimmers at low Reynolds numbers, e .g. the Copepod swimmer (an abundant variety of zooplankton) or the Purcell swimmer. The problem was studied from the point of view of geometric optimal control [17], [18] combining theoretical and numerical computations or controllability techniques [19].

6.3. Averaging in control and application to space mechanics

Participants: Bernard Bonnard, Jean-Baptiste Caillau, Helen-Clare Henninger, Jana Němcová [Institute of Chemical Tech, Prague, CZ], Jean-Baptiste Pomet, Jeremy Rouot.

We have obtained results on the structure of the average system for the planar minimum time problem without perturbation in [4], and the “double average” that takes the lunar perturbation into account in [13]. This is also the topic of Helen Henninger’s PhD [1].

The structure of the problem where the consumption (i.e. the L^1 norm of the control) is minimized is studied in [5].

The book [16] is a general reference opus edited by members of the team.

MEMPHIS Team

7. New Results

7.1. Plastic impact of iron on aluminium

A new model for plasticity has been developed this year. An iron projectile is impacting an aluminium plate immersed in air. The initial horizontal velocity of the iron is $1000m.s^{-1}$. The computation is performed on a 2000×1600 mesh with 144 processors. Homogeneous Neumann conditions are imposed on the left and right borders and embedded on the top and bottom.

The results are presented in Fig 6 with a schlieren image (bottom) and the von Mises criteria (top) at different time steps. A log scale is used and the minimum value is fixed to 10^9 . We can see that the plate is strongly deformed and form at the end a filament. The projectile is flattened but not as much as in the literature because the yield plastic limit is higher. We see a longitudinal wave propagating in the plate followed by a shear wave that causes the plasticity of the material.

7.2. Air-helium shock-bubble interaction

A three-dimensional hyperelastic model has been developed. It can deal with multi-fluid and solids. Here we show the propagation in air of a Mach 1.22 shock through an helium bubble. The computation is performed on a $1000 \times 400 \times 400$ mesh and lasts for 50h on 300 processors. The zero iso-value of the level set function and schlieren on the horizontal plane through the center of the bubble are presented at different times on Fig. 7 .

7.3. Particles flowing in a fluid

A new type of algorithm is designed to enable contacts efficiently between particles immersed in a fluid by adding a short range repulsive force. The algorithm is derived from the multi geometric deformable model introduced for image segmentation. It can handle multiple deforming bodies and avoid collision using a short range repulsive force depending on the distance to the closest interface. The main advantages of this method is it requires only five fields (three label maps and two distance functions) and one level set function to capture an arbitrary number of cells and it can, at the same time, deal with collisions.

7.4. Inertial Sea Wave Energy Converter (ISWEC)

The ISWEC is a floater and was design by Wave For Energy (<http://www.waveforenergy.com>) to extract the energy of typical waves in the Mediterranean Sea. The energy is extracted using a mechanical system based on a gyroscope activated by the motion of the floater generated by sea waves. This is a complex system coupling Fluid/Fluid/structure interfaces, computation of the rigid motion of the floater and computation of the power extraction. The problem is solved using in-house numerical solver (NaSCar) developed in MEMPHIS team. The interfaces are tracked using level se functions. The bi fluid interface is computed using Continuous Surface Force method (CSF), the motion of the floater imposed by penalization is computed using the forces and the torques exerted by the flow, and finally this motion activates the gyroscope for power extraction. The gyroscope model was developed by the Politecnico di Torino. Figure 1 shows a numerical simulation of the iswec (see <http://www.math.u-bordeaux1.fr/mbergman/> for a movie)

7.5. Flow with many particles

A version of the code NaSCar has been developed to simulated the flow around particules with high volume fraction (see figure 10). The standard central lubrication forces are used to computed the interaction between spherical particules. Ongoing project will deal with non spherical particules (Lisl Weynans and PhD Baptiste Lambert).

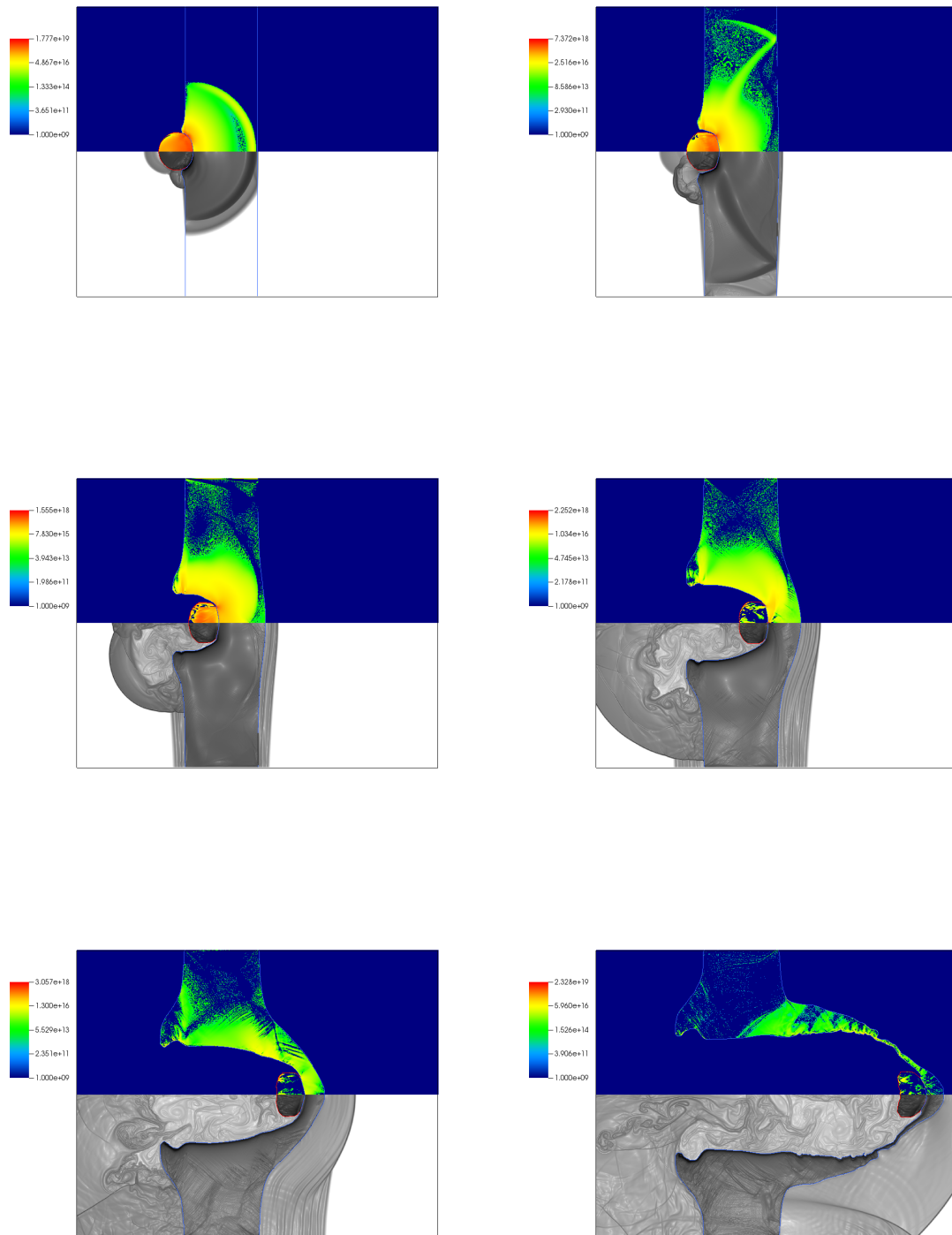


Figure 6. Impact of Iron on Aluminium TC2. Schlieren image and von Mises criteria at $t = 0.03\text{ms}$, 0.06ms , 0.13ms , 0.26ms , 0.53ms and $t = 1.04\text{ms}$ from left to right, top to bottom

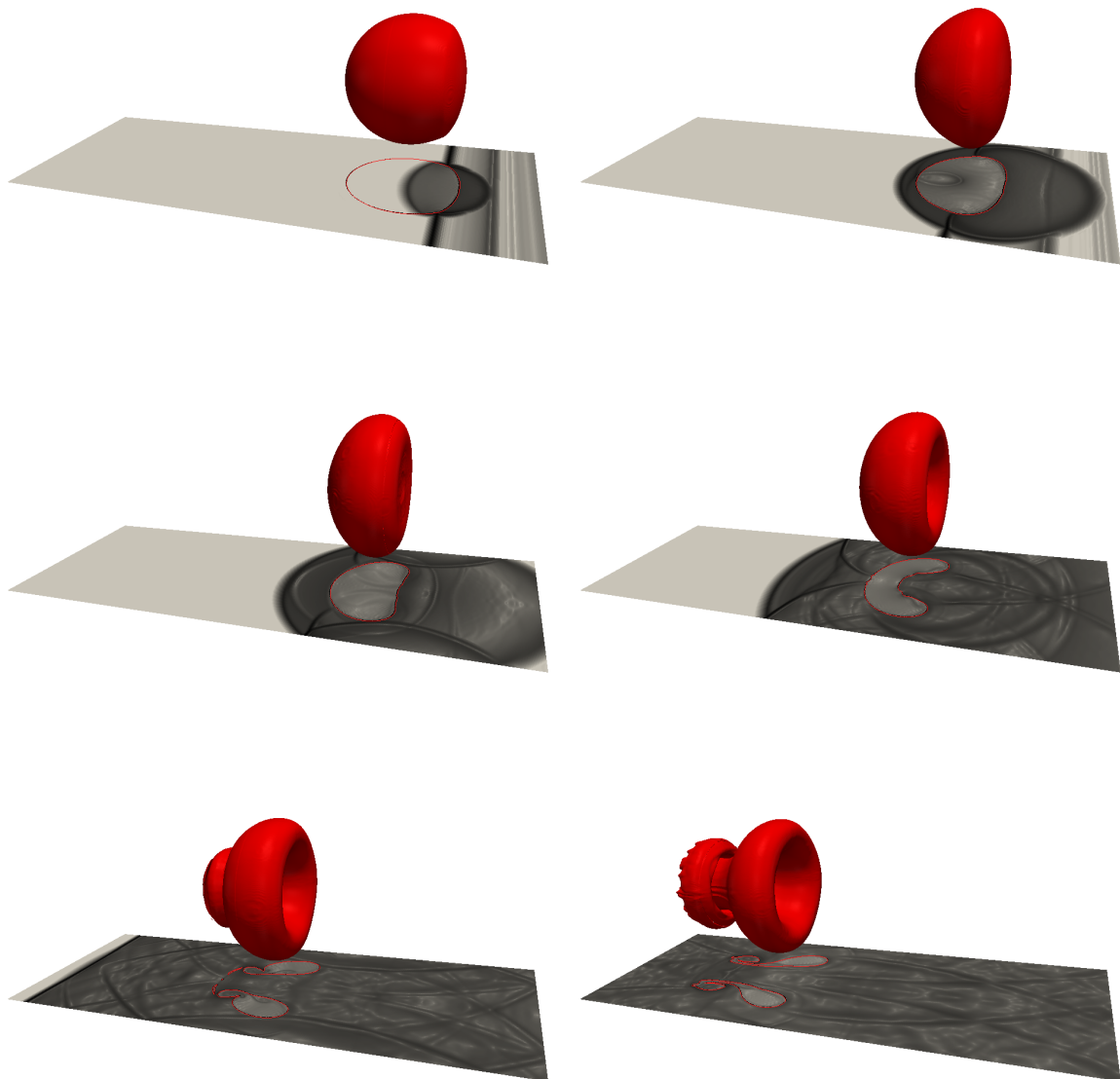


Figure 7. Interaction of a Mach 1.22 shock propagating in air through an helium bubble (TC1). Pictures at $t = 62\mu s, 110\mu s, 163\mu s, 264\mu s, 471\mu s, 735\mu s$. From left to right, top to bottom.

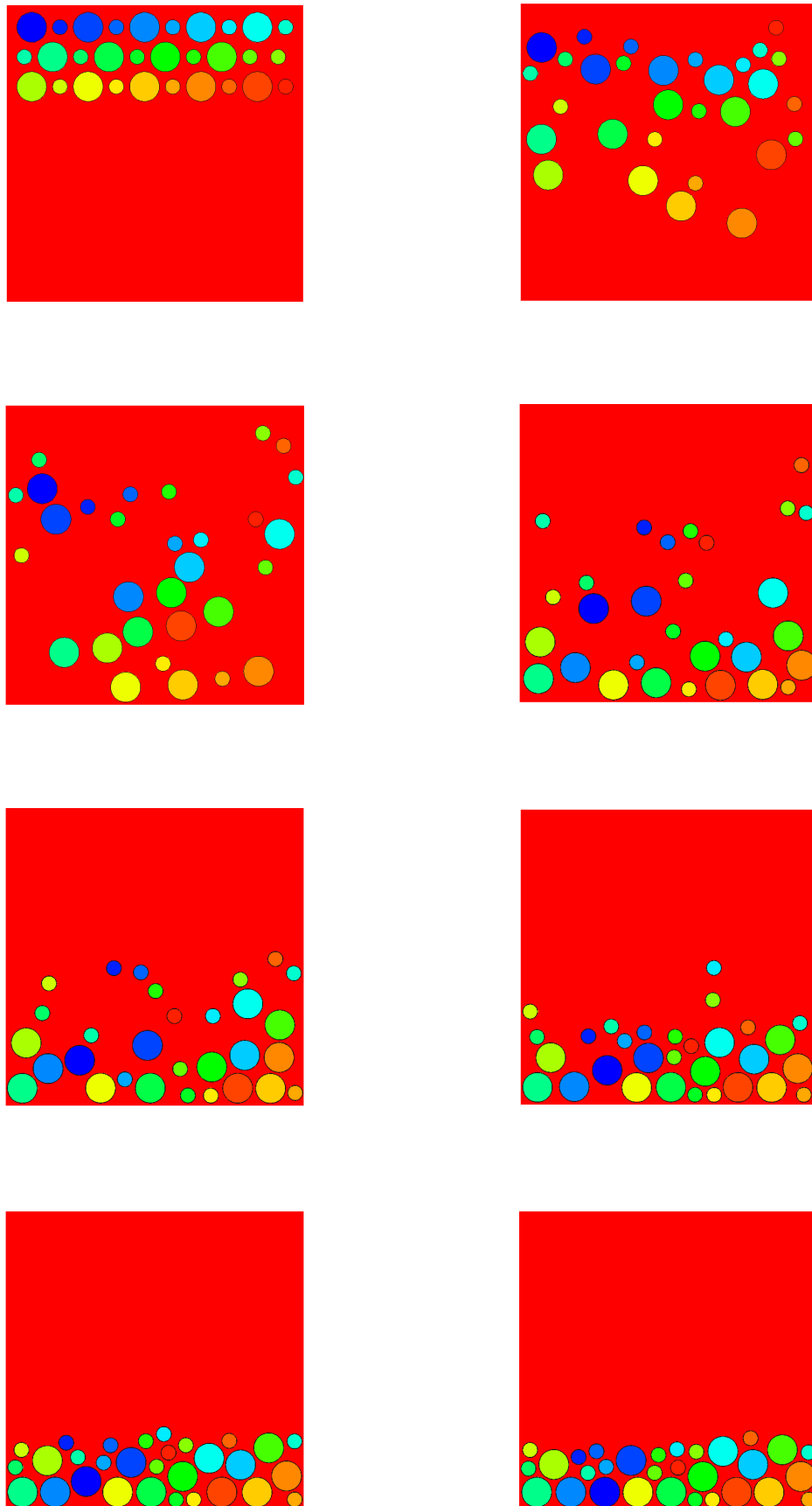


Figure 8. Simulation of 30 rigid bodies of different radii ($R = 0.05$ or $R = 0.025$) falling under gravity. The colors indicate the values of the first label map from dark blue for the first body to dark orange for the 30th body and red for the fluid that is the 31st object.

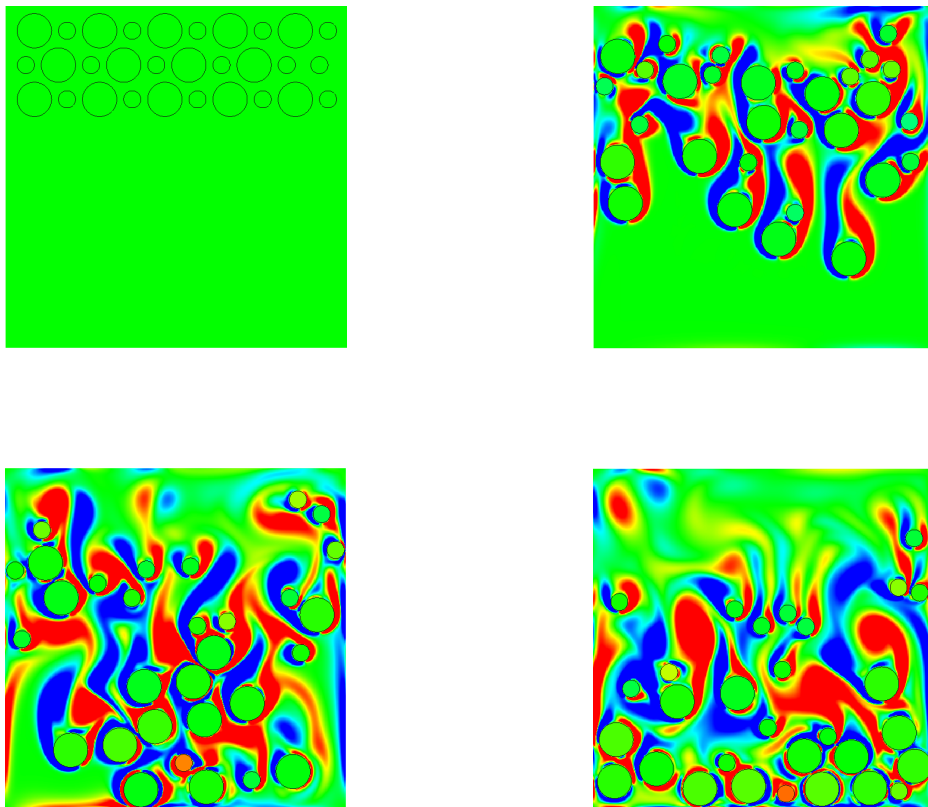


Figure 9. Simulation of 30 rigid bodies of different radii ($R = 0.05$ or $R = 0.025$) falling under gravity. The colors indicate the vorticity levels from dark blue for -200 and dark red for 200.

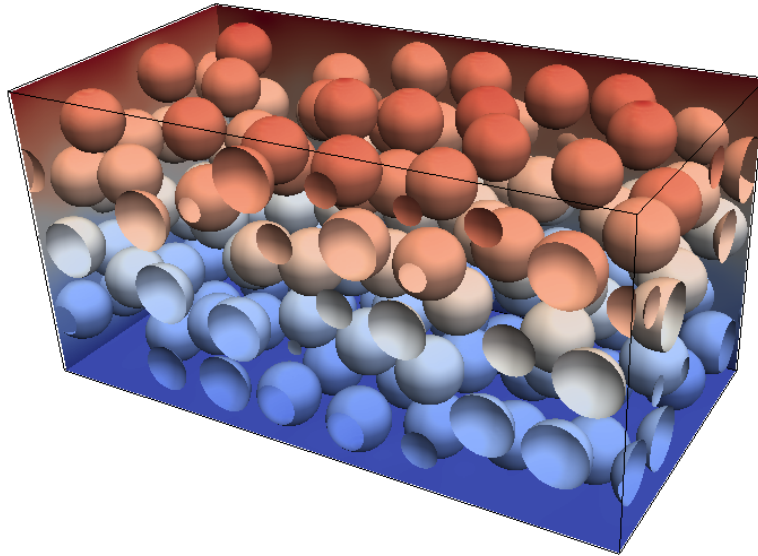


Figure 10. Rheology of a Couette flow with solid particles.

7.6. Overset method

With Valeol (Cifre PhD Claire Morel) we are developing an Overlapping grids approach coupling a background cartesian grid with a body fitted grid around wind turbine blades. This method allows us to push away the limit of the numerical simulation on octree grids when small boundary layers play an important role. The generation of overset grids is based on level set functions (post doc AMIES by Franck Luddens). This method has been implemented for two dimensional test cases using the Schwartz method to deal with the domain decomposition. In the same time, we are also building the global operators onto the whole domain, i.e. without Schwartz iterations (PhD Federico Tesser).

7.7. Electrostrictive materials: modelling and simulation

In this work, a result of the collaboration between physicists, chemists (Annie Colin and Philippe Poulin, CRPP Bordeaux) and applied mathematicians, we deal with mathematical modelling and simulation of electrostrictive materials. These kinds of materials are composed by a polymeric matrix with carbon nanotubes embedded in and this structure gives them interesting electrical properties. Their dielectric constant varies as a function of the mechanical deformation. Housed in a capacitor, they show variable capacity when subjected to vibration and they can generate potential differences from mechanical deformations. Because of their composition, their structure involves different physical scales, from the small nanotube dimension, through the scale of nanotube clusters, to the large dimension of the sample. Our purpose is to provide physicists and chemists with a tool to test *in silico* several material configurations and to have a deeper insight into the features of these materials, developing numerical models which can predict their steady and unsteady behaviour. We propose to model the physical problem by reducing the nanotubes to dipoles and solving a Gauss equation for the electrical potential equation informed of the presence of nanotubes with zero electrical field conditions on the centres of the nanotubes. We started considering the steady problem, which is interesting for the purpose

of understanding the basic electrical properties of different nanotube configurations and of designing of the material. In order to discretize and simulate the mathematical problem, we chose parallel linear octree-based adaptive meshes and we developed an original hybrid Finite Volume/Finite Difference second-order scheme for 2D and 3D elliptic problems on this kind of mesh. A convergence analysis of the numerical scheme has been developed and validating test cases have been performed. Good qualitative agreement between numerical and real experiments has been observed for the steady model. In the future we aim to quantitatively compare the numerical results and the real material behaviour, to model the unsteady problem and to deal with electrical consequences of mechanical deformations.

7.8. Development of a sharp cartesian method for the simulation of flows with high density ratios

We have developed a sharp cartesian method for the simulation of incompressible flows with high density and viscosity ratios, like air-water interfaces. This method is inspired from a second-order cartesian method for elliptic problems with immersed interfaces (Cisternino-Weynans 2012). A classical predictor-corrector algorithm is used to solve the fluid equations, in a non-incremental version, which means that the guess value for the pressure is zero. This choice avoids instability issues due to the discontinuous pressure values when the interface moves. We take into account the viscous forces by regularizing the density and viscosity values. This approach allows for a more straightforward and robust treatment, and has been proven to provide satisfactory accuracies for high Reynolds numbers. To compute the pressure, it is necessary to solve an elliptic problem. This elliptic problem with discontinuous values across an interface is solved with the second order method cited above. The originality of this method lies in the use of unknowns located at the interface. These interface unknowns are used to discretize the flux jump conditions and the elliptic operator accurately enough to get a second order convergence in maximum norm.

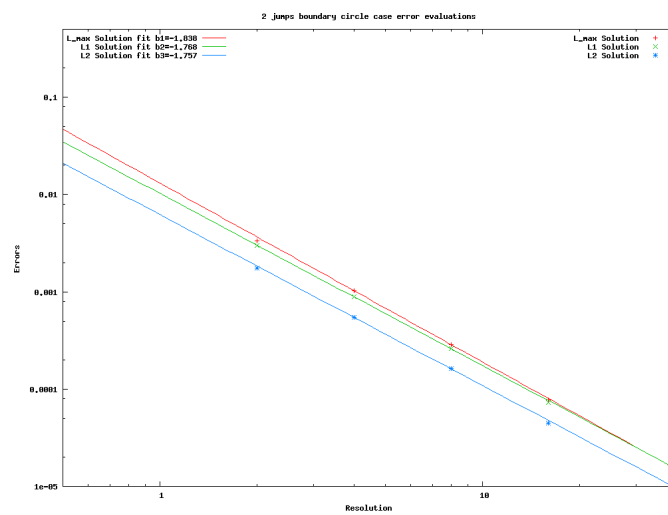
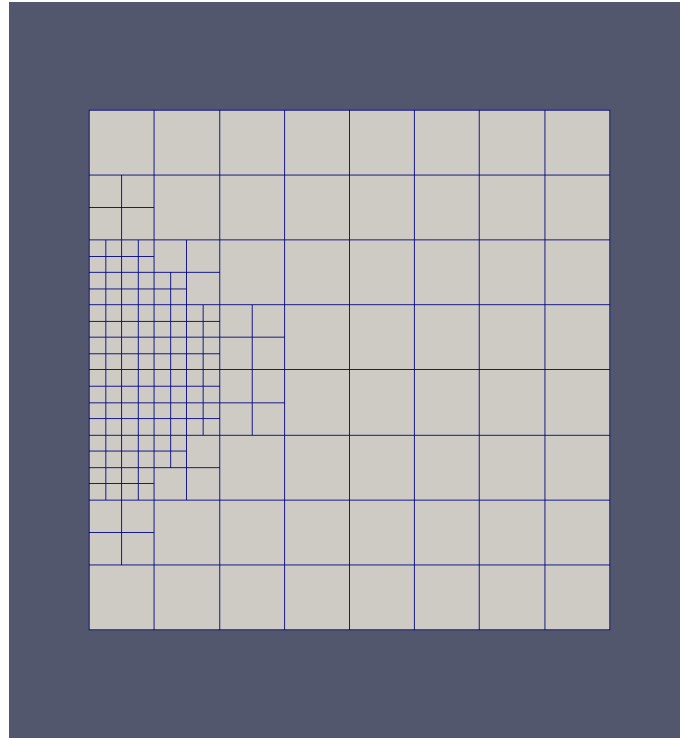


Figure 11. Electrostrictive materials: Mesh and convergence results for the solver.

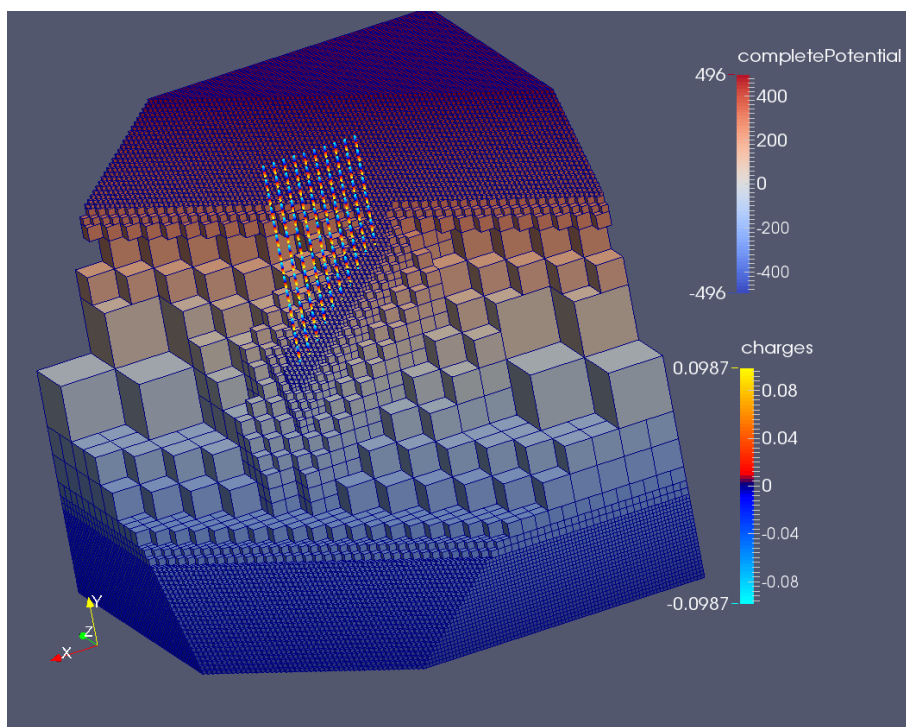


Figure 12. Electrostrictive materials: 3D configuration with nanotubes

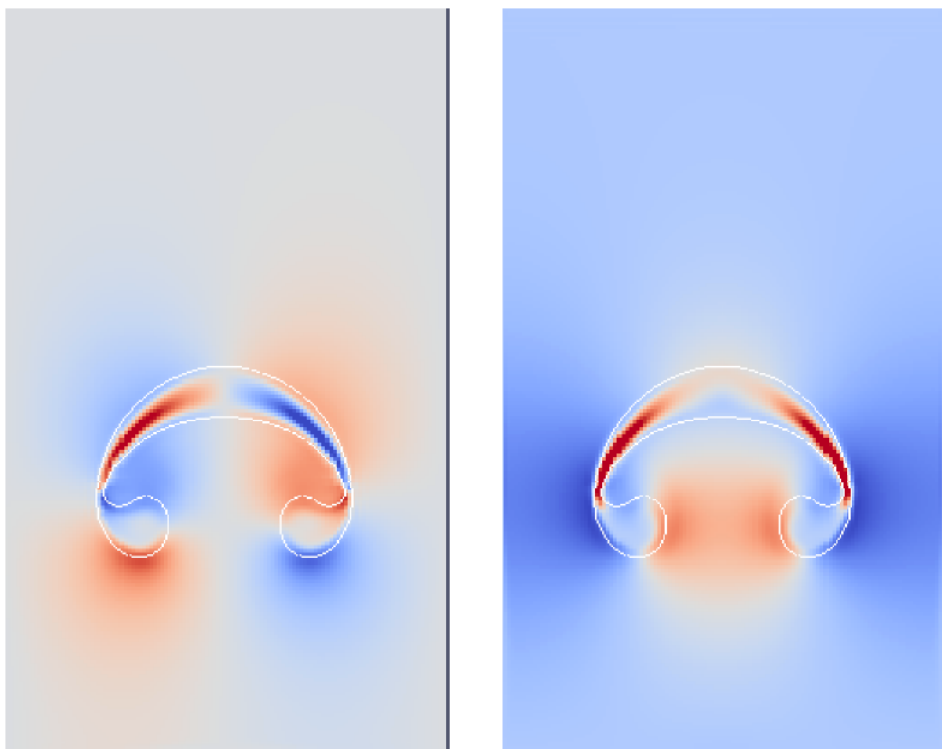


Figure 13. Horizontal et vertical velocity profile of an air bubble rising into water.

MEPHYSTO Team

7. New Results

7.1. Quantitative stochastic homogenization

7.1.1. Discrete equations

7.1.1.1. Decay of the semi-group

A. Gloria, S. Neukamm (Univ. Dresden), and F. Otto (MPI for mathematics in the sciences, Leipzig) developed in [15] a general approach to quantify ergodicity in stochastic homogenization of scalar discrete elliptic equations. Using a parabolic approach, they obtained optimal estimates on the time-decay of the so-called environment seen from the particle. This allowed them to prove optimal bounds on the corrector gradient and the corrector itself in any dimension (thus improving on [5]). They also obtained the first error analysis of the popular periodization method to approximate the homogenized coefficients.

7.1.1.2. Quantitative CLT

In [16], A. Gloria and J. Nolen (Duke Univ.) proved a quantitative central limit theorem for the effective conductance on the discrete torus. In particular, they quantified the Wasserstein distance between a normal random variable and the CLT-like rescaling of the difference between the approximation of the effective conductance by periodization and the effective conductance. Their estimate is sharp and shows that the Wasserstein distance goes to zero (up to logarithmic factors) as if the energy density of the corrector was iid (which it is not). This completes and settles the analysis started in [15] on the approximation of homogenized coefficients by periodization by characterizing the limiting law in addition to the scaling.

7.1.2. Continuum equations

7.1.2.1. Scalar equations with random coefficients

In [17], A. Gloria and F. Otto extended their results [4], [5] on discrete elliptic equations to the continuum setting. They treated in addition the case of non-symmetric coefficients, and obtained optimal estimates in all dimensions by the elliptic approach (whereas [4],[5] were suboptimal for $d = 2$).

In [14], A. Gloria and D. Marahrens (MPI for mathematics in the sciences, Leipzig) extended the annealed results [14] on the discrete Green function by D. Marahrens and F. Otto to the continuum setting. As a by-product of their result, they obtained new results in uncertainty quantification by estimating optimally the variance of the solution of an elliptic PDE whose coefficients are perturbed by some noise with short range of dependence.

7.1.2.2. Systems with random coefficients

In a revised version of [58], A. Gloria, S. Neukamm, and F. Otto developed a regularity theory for random elliptic operators inspired by the contributions of Avellaneda and Lin [43] in the periodic setting and of S. Armstrong with C. Smart [42]. This allowed them to consider coefficients with arbitrarily slow decaying correlations in the form of a family of correlated Gaussian fields, and obtain (in the new version of this paper) a family of estimates with optimal rates and exponential-type integrability.

In [35], A. Gloria and F. Otto obtained the first nearly-optimal estimates with optimal stochastic integrability on the corrector for linear elliptic systems whose coefficients satisfy a finite range of dependence assumption (thus avoiding the functional inequalities they considered so far).

7.1.2.3. Systems with almost periodic coefficients

In [23], S. Armstrong, A. Gloria and T. Kuusi (Aalto University) obtained the first improvement over the thirty year-old result by Kozlov [60] on almost periodic homogenization. In particular they introduced a class of almost periodic coefficients which are not quasi-periodic (and thus strictly contains the Kozlov class) and for which almost periodic correctors exist. Their approach combines the regularity theory developed by S. Armstrong and C. Smart in [42] and adapted to the almost periodic setting by S. Armstrong and Z. Shen [41], a new quantification of almost-periodicity, and a sensitivity calculus in the spirit of [4].

7.1.3. Clausius-Mossotti formulas

In the mid-nineteenth century, Clausius, Mossotti and Maxwell essentially gave a first order Taylor expansion for (what is now understood as) the homogenized coefficients associated with a constant background medium perturbed by diluted spherical inclusions. Such an approach was recently used and extended by the team MATERIALS to reduce the variance in numerical approximations of the homogenized coefficients, cf. [39], [38], [62]. In [12], M. Duerinckx and A. Gloria gave the first rigorous proof of the Clausius-Mossotti formula and provided the theoretical background to analyze the methods introduced in [62].

7.2. Derivation of nonlinear elasticity from polymer-physics

7.2.1. Reconstruction of analytical constitutive laws

In [10], M. de Buhan (CNRS, Univ. Paris Descartes), A. Gloria, P. Le Tallec and M. Vidrascu proposed a numerical method to produce analytical approximations (that can be used in practical nonlinear elasticity softwares) of the numerical approximations obtained in [57] of the discrete-to-continuum energy density derived theoretically in [1]. This numerical method is based on the parametrization of the set of polyconvex Ogden laws and on the combination of a least square method and a genetic algorithm (cf. CMA-ES, <https://www.lri.fr/hansen/cmaesintro.html>).

7.2.2. Stochastic homogenization of unbounded integral functionals

In [34], M. Duerinckx and A. Gloria succeeded in relaxing one of the two unphysical assumptions made in [1] on the growth of the energy of polymer chains. In particular, [34] deals with the case when the energy of the polymer chain is allowed to blow up at finite deformation.

7.3. Numerical methods

7.3.1. Numerical homogenization

Inspired by the quantitative analysis of [15] and [17], Z. Habibi (former SIMPAF post-doctoral fellow) and A. Gloria introduced in [13] a general method to reduce the so-called resonance error in numerical homogenization, both at the levels of the approximation of the homogenized coefficients and of the correctors. This method significantly extends [2]. The method relies on the introduction of a massive term in the corrector equation and of a systematic use of Richardson extrapolation. In the three academic examples of heterogeneous coefficients (periodic, quasiperiodic, and Poisson random inclusions), the method yields optimal theoretical and empirical convergence rates, and outperforms most of the other existing methods.

7.3.2. Numerical methods for evolution equations

In [11], G. Dujardin and P. Lafitte (ECP) published a result on the asymptotic behavior of splitting schemes applied to multiscale systems which have strongly attracting equilibrium states. They proposed a definition of the asymptotic order of such schemes and proved on examples of ODEs and PDEs systems that one can achieve high asymptotic order with such schemes, provided sufficient conditions are fulfilled.

In [25], G. Dujardin proposed to use high order methods for the numerical simulation of rotating Bose-Einstein condensates. With his co-authors, he developed exponential Runge-Kutta methods and Lawson method for this problem and he analyzed the convergence order of these methods. In particular, they proved that one can achieve maximal order $2s$ with methods with s -stages. They also supported their analysis with numerical experiments carried out in physically realistic simulations.

7.4. Schrödinger equations

7.4.1. Nonlinear optic fibers

In [18], S. De Bièvre, G. Dujardin, and S. Rota-Noradi, in collaboration with physicists of the PhLAM laboratory in Lille, developed an analysis of the phenomenon of modulational instability in dispersion-kicked optical fibers. They proposed a genuine analysis of the phenomenon, together with estimates on physical properties such as the gain along the fibers, and they showed that their analysis actually fits both numerical and physical experiments.

In [20], S. De Bièvre and G. Dujardin, in collaboration with physicists of the PhLAM laboratory in Lille, developed an analysis of the propagation along a periodically-modulated optic fiber of generalized Peregrine rogue waves. In particular, they provided a full analysis of the multiple compression points appearing in such waves.

In D. Bonheure and R. Nascimento [21] obtained new results on the existence and qualitative properties of waveguides for a mixed-diffusion NLS. They provided a full qualitative description of the waveguides when the fourth order dissipation is small.

7.4.2. Nonlinear Schrödinger equations

S. De Bièvre, S. Rota Nodari, and F. Genoud (CEMPI visitor, September 2013) have explained the geometry underlying the so-called energy-momentum method for proving orbital stability in infinite dimensional Hamiltonian systems. Applications include the orbital stability of solitons of the NLS and Manakov equations. This work appeared as a chapter (120p) in the first volume of the CEMPI Lecture Notes in Mathematics, cf. [48].

In [26], Bonheure, S. Cingolani and M. Nys obtained new striking results on stationary solutions of the 3D NLS driven by an exterior magnetic field. They construct a new class of cylindrical solutions in the energy class which concentrate, in the semi-classical limit, on a circle of the plane through the equator. In contrast with the case of solutions localized around a single point, the concentration is driven by the electrical field as well as the magnetic field

7.5. Stochastic acceleration and approach to equilibrium

S. De Bièvre and E. Soret rigorously proved the growth rate of the energy in a Markovian model for stochastic acceleration of a particle in a random medium, cf. [67] and [7].

S. De Bièvre, Carlos Mejia-Monasterio (Madrid) and Paul E. Parris (Missouri) [49] studied thermal equilibration in a two-component Lorentz gas, in which the obstacles are modeled by rotating disks. They show that a mechanism of dynamical friction leads to a fluctuation-dissipation relation that is responsible for driving the system to equilibrium.

Stephan De Bièvre, Jeremy Faupin (Metz) and Schuble (Metz) [32] studied a related model quantum mechanically. Here a quantum particle moves through a field of quantized bose fields, modeling membranes that exchange energy and momentum with the particle. They establish a number of spectral properties of this model, that will be essential to study the time-asymptotic behavior of the system.

7.6. Miscellaneous results

In [24] A. Benoit proved that for linear hyperbolic systems of equations in the quarter space a violent instability can be caused by the accumulation of an arbitrary large number of weak instabilities. The proof of this result is based on the construction of the WKB expansions for hyperbolic corner problems with self-interacting phases and is a continuation of [45].

In [9], C. Cancès, T. Gallouët, and L. Monsaingeon gave a gradient flow interpretation for incompressible immiscible two-phase flows in porous media. With C. Chainais-Hillairet, T. Gallouët characterized the pseudo-stationary state for a corrosion model in [31].

In [8], D. Bonheure, E. Moreira dos Santos, M. Ramos and H. Tavares construct least energy nodal solutions of Hamiltonian elliptic systems. The construct is tricky since the functional associated to Hamiltonian elliptic systems is strongly indefinite. The proof uses a dual variational argument and an approximation scheme with some ideas of Gama-convergence type.

In [27], D. Bonheure, P. D'Avenia and A. Pomponio aim to derive rigorously the PDE formulation of the Born-Infeld model in the electrostatic case. This nonlinear model of electromagnetism was introduced by Born and Infeld who proposed a new Lagrangian which theoretically assumes the existence of a maximal field intensity, likewise Einstein's Lagrangian of special relativity opposed to Newton's Lagrangian of classical mechanics. The paper contains new results and new insights on the model. It covers several relevant particular cases but we are still far from the full understanding of the problem.

In [29] and [30], D. Bonheure and coauthors study patterns and phase transitions in a fourth order extension of the famous Allen-Cahn model. In [29], some rigidity results à la Gibbons are proved while [30] concerns qualitative properties of positive patterns with Navier boundary conditions. A conjecture related to De Giorgi's famous one concerning the one dimensionality of monotone phase transition in the classical Allen Cahn model is proposed in [29].

In [28], D. Bonheure and collaborators study multi-layer solutions of the Lin-Ni-Takagi model, which comes from the Keller-Segel model of chemotaxis in a specific case. A remarkable feature of the results is that the layers do not accumulate to the boundary of the domain but satisfy an optimal partition problem contrary to the previous type of solutions constructed for this model.

In [53], M. Duerinckx proved a new mean-field limit result for the gradient flow evolution of particle systems with pairwise Riesz interactions, in dimensions 1 and 2, in cases for which this problem was still open. The proof is based on a method introduced by Serfaty [66] in the context of the Ginzburg-Landau vortices, using regularity and stability properties of the limiting equation.

MISTIS Project-Team

7. New Results

7.1. Mixture models

7.1.1. Taking into account the curse of dimensionality

Participants: Stéphane Girard, Alessandro Chiancone, Seydou-Nourou Sylla.

Joint work with: C. Bouveyron (Univ. Paris 5), M. Fauvel (ENSAT Toulouse) and J. Chaussois (Gipsa-lab and Grenoble-INP)

In the PhD work of Charles Bouveyron (co-advised by Cordelia Schmid from the Inria LEAR team) [61], we proposed new Gaussian models of high dimensional data for classification purposes. We assume that the data live in several groups located in subspaces of lower dimensions. Two different strategies arise:

- the introduction in the model of a dimension reduction constraint for each group
- the use of parsimonious models obtained by imposing to different groups to share the same values of some parameters

This modelling yields a supervised classification method called High Dimensional Discriminant Analysis (HDDA) [4]. Some versions of this method have been tested on the supervised classification of objects in images. This approach has been adapted to the unsupervised classification framework, and the related method is named High Dimensional Data Clustering (HDDC) [3]. Our recent work consists in adding a kernel in the previous methods to deal with nonlinear data classification and heterogeneous data [13]. We first investigate the use of kernels derived from similarity measures on binary data [30]. The targeted application is the analysis of verbal autopsy data (PhD thesis of N. Sylla): Indeed, health monitoring and evaluation make more and more use of data on causes of death from verbal autopsies in countries which do not keep records of civil status or with incomplete records. The application of verbal autopsy method allows to discover probable cause of death. Verbal autopsy has become the main source of information on causes of death in these populations. Second, the kernel classification method is applied to three real hyperspectral data sets, and compared with three other classifiers. The proposed models show good results in terms of classification accuracy and processing time [21].

7.2. Semi and non-parametric methods

7.2.1. Conditional extremal events

Participant: Stéphane Girard.

Joint work with: L. Gardes (Univ. Strasbourg), G. Mazo (Univ. Catholique de Louvain), J. El Methni (Univ. Paris 5) and S. Louhichi (Univ. Grenoble 1)

The goal of the PhD theses of Alexandre Lekina and Jonathan El Methni was to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, *ie* the situation where some covariate information X is recorded simultaneously with a quantity of interest Y . In such a case, the tail heaviness of Y depends on X , and thus the tail index as well as the extreme quantiles are also functions of the covariate. We combine nonparametric smoothing techniques [63] with extreme-value methods in order to obtain efficient estimators of the conditional tail index and conditional extreme quantiles. The strong consistency of such an estimator is established in [53]. When the covariate is functional and random (random design) we focus on kernel methods [23].

Conditional extremes are studied in climatology where one is interested in how climate change over years might affect extreme temperatures or rainfalls. In this case, the covariate is univariate (time). Bivariate examples include the study of extreme rainfalls as a function of the geographical location. The application part of the study is joint work with the LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) located in Grenoble [20] and the "département Génie urbain" of "Université Paris-Est Marne-la-vallée" [11].

7.2.2. Estimation of extreme risk measures

Participant: Stéphane Girard.

Joint work with: A. Daouia (Univ. Toulouse), E. Deme (Univ. Gaston-Berger, Sénégal), A. Guillou (Univ. Strasbourg) and G. Stupfler (Univ. Aix-Marseille).

One of the most popular risk measures is the Value-at-Risk (VaR) introduced in the 1990's. In statistical terms, the VaR at level $\alpha \in (0, 1)$ corresponds to the upper α -quantile of the loss distribution. The Value-at-Risk however suffers from several weaknesses. First, it provides us only with a pointwise information: $\text{VaR}(\alpha)$ does not take into consideration what the loss will be beyond this quantile. Second, random loss variables with light-tailed distributions or heavy-tailed distributions may have the same Value-at-Risk. Finally, Value-at-Risk is not a coherent risk measure since it is not subadditive in general. A first coherent alternative risk measure is the Conditional Tail Expectation (CTE), also known as Tail-Value-at-Risk, Tail Conditional Expectation or Expected Shortfall in case of a continuous loss distribution. The CTE is defined as the expected loss given that the loss lies above the upper α -quantile of the loss distribution. This risk measure thus takes into account the whole information contained in the upper tail of the distribution. It is frequently encountered in financial investment or in the insurance industry. In [51], we have established the asymptotic properties of the CTE estimator in case of extreme losses, *i.e.* when $\alpha \rightarrow 0$ as the sample size increases. We have exhibited the asymptotic bias of this estimator, and proposed a bias correction based on extreme-value techniques. A second possible coherent alternative risk measure is based on expectiles [59]. Compared to quantiles, the family of expectiles is based on squared rather than absolute error loss minimization. The flexibility and virtues of these least squares analogues of quantiles are now well established in actuarial science, econometrics and statistical finance. Both quantiles and expectiles were embedded in the more general class of M-quantiles as the minimizers of a generic asymmetric convex loss function. It has been proved very recently that the only M-quantiles that are coherent risk measures are the expectiles.

7.2.3. Multivariate extremal events

Participants: Stéphane Girard, Florence Forbes.

Joint work with: F. Durante (Univ. Bolzen-Bolzano, Italy) L. Gardes (Univ. Strasbourg) and G. Mazo (Univ. Catholique de Louvain, Belgique).

Copulas are a useful tool to model multivariate distributions [67].

However, while there exist various families of bivariate copulas, much fewer has been done when the dimension is higher. To this aim an interesting class of copulas based on products of transformed copulas has been proposed in the literature. The use of this class for practical high dimensional problems remains challenging. Constraints on the parameters and the product form render inference, and in particular the likelihood computation, difficult. We proposed a new class of high dimensional copulas based on a product of transformed bivariate copulas [26]. No constraints on the parameters refrain the applicability of the proposed class which is well suited for applications in high dimension. Furthermore the analytic forms of the copulas within this class allow to associate a natural graphical structure which helps to visualize the dependencies and to compute the likelihood efficiently even in high dimension. The extreme properties of the copulas are also derived and an R package has been developed.

As an alternative, we also proposed a new class of copulas constructed by introducing a latent factor. Conditional independence with respect to this factor and the use of a nonparametric class of bivariate copulas lead to interesting properties like explicitness, flexibility and parsimony. In particular, various tail behaviours are exhibited, making possible the modeling of various extreme situations [19], [27], [52]. A pairwise moment-based inference procedure has also been proposed and the asymptotic normality of the corresponding estimator has been established [28].

In collaboration with L. Gardes, we investigate the estimation of the tail copula, which is widely used to describe the amount of extremal dependence of a multivariate distribution. In some situations such as risk management, the dependence structure can be linked with some covariate. The tail copula thus depends on this covariate and is referred to as the conditional tail copula. The aim of our work is to propose a nonparametric estimator of the conditional tail copula and to establish its asymptotic normality [22].

7.2.4. Level sets estimation

Participant: Stéphane Girard.

Joint work with: G. Stupfler (Univ. Aix-Marseille)

The boundary bounding the set of points is viewed as the larger level set of the points distribution. This is then an extreme quantile curve estimation problem. We proposed estimators based on projection as well as on kernel regression methods applied on the extreme values set, for particular set of points [10]. We also investigate the asymptotic properties of existing estimators when used in extreme situations. For instance, we have established in collaboration with G. Stupfler that the so-called geometric quantiles have very counter-intuitive properties in such situations [24], [25] and thus should not be used to detect outliers.

7.2.5. Retrieval of Mars surface physical properties from OMEGA hyperspectral images.

Participants: Stéphane Girard, Alessandro Chiancone.

Joint work with: J. Chanussot (Gipsa-lab and Grenoble-INP).

Visible and near infrared imaging spectroscopy is one of the key techniques to detect, to map and to characterize mineral and volatile (eg. water-ice) species existing at the surface of planets. Indeed the chemical composition, granularity, texture, physical state, etc. of the materials determine the existence and morphology of the absorption bands. The resulting spectra contain therefore very useful information. Current imaging spectrometers provide data organized as three dimensional hyperspectral images: two spatial dimensions and one spectral dimension. Our goal is to estimate the functional relationship F between some observed spectra and some physical parameters. To this end, a database of synthetic spectra is generated by a physical radiative transfer model and used to estimate F . The high dimension of spectra is reduced by Gaussian regularized sliced inverse regression (GRSIR) to overcome the curse of dimensionality and consequently the sensitivity of the inversion to noise.

In his PhD thesis work, Alessandro Chiancone studies the extension of the SIR method to different sub-populations. The idea is to assume that the dimension reduction subspace may not be the same for different clusters of the data [14].

7.2.6. Robust Sliced Inverse Regression.

Participants: Stéphane Girard, Alessandro Chiancone, Florence Forbes.

Sliced Inverse Regression (SIR) has been extensively used to reduce the dimension of the predictor space before performing regression. Recently it has been shown that this techniques is, not surprisingly, sensitive to noise. Different approaches has been proposed to robustify SIR, in this work, we start considering an inverse problem proposed by R.D. Cook and we show that the framework can be extended to take into account a non-Gaussian noise. Generalized Student distribution are considered and all parameters are estimated via EM algorithm. The algorithm is outlined and tested comparing the results with different approaches on simulated data. Results on a real dataset shows the interest of this technique in presence of outliers.

7.2.7. Robust Locally linear mapping with mixtures of Student distributions

Participants: Florence Forbes, Emeline Perthame, Brice Olivier, Leo Nicoletti.

The standard GLLiM model [17] for high dimensional regression assumes Gaussian noise models and is in its unconstrained version equivalent to a joint GMM. The fact that response and independent variables (X, Y) are jointly a mixture of Gaussian distribution is the key for all derivations in the model. In this work, we show that similar developments are possible based on a joint Student Mixture model, joint SMM. It follows a new model referred to as SLLiM for Student Locally linear mapping for which we investigate the robustness to outlying data in a high dimensional regression context.

7.3. Markov models

7.3.1. Change-point models for tree-structured data

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Inria) and Yann Guédon (CIRAD), Inria Virtual Plants.

In the context of plant growth modelling, methods to identify subtrees of a tree or forest with similar attributes have been developed. They rely either on hidden Markov modelling or multiple change-point approaches. The latter are well-developed in the context of sequence analysis, but their extensions to tree-structured data are not straightforward. Their advantage on hidden Markov models is to relax the strong constraints regarding dependencies induced by parametric distributions and local parent-children dependencies. Heuristic approaches for change-point detection in trees were proposed and applied to the analysis of patchiness patterns (consisting of canopies made of clumps of either vegetative or flowering botanical units) in mango trees [45].

7.3.2. Hidden Markov models for the analysis of eye movements

Participants: Jean-Baptiste Durand, Brice Olivier.

This research theme is supported by a LabEx PERSYVAL-Lab project-team grant.

Joint work with: Marianne Clausel (LJK) Anne Guérin-Dugué (GIPSA-lab) and Benoit Lemaire (Laboratoire de Psychologie et Neurocognition)

In the last years, GIPSA-lab has developed computational models of information search in web-like materials, using data from both eye-tracking and electroencephalograms (EEGs). These data were obtained from experiments, in which subjects had to make some kinds of press reviews. In such tasks, reading process and decision making are closely related. Statistical analysis of such data aims at deciphering underlying dependency structures in these processes. Hidden Markov models (HMMs) have been used on eye movement series to infer phases in the reading process that can be interpreted as steps in the cognitive processes leading to decision. In HMMs, each phase is associated with a state of the Markov chain. The states are observed indirectly through eye-movements. Our approach was inspired by Simola *et al.* (2008) [68], but we used hidden semi-Markov models for better characterization of phase length distributions. The estimated HMM highlighted contrasted reading strategies (i.e., state transitions), with both individual and document-related variability.

However, the characteristics of eye movements within each phase tended to be poorly discriminated. As a result, high uncertainty in the phase changes arose, and it could be difficult to relate phases to known patterns in EEGs.

This is why, as part of Brice Olivier's PhD thesis, we are developing integrated models coupling EEG and eye movements within one single HMM for better identification of the phases. Here, the coupling should incorporate some delay between the transitions in both (EEG and eye-movement) chains, since EEG patterns associated to cognitive processes occur lately with respect to eye-movement phases. Moreover, EEGs and scanpaths were recorded with different time resolutions, so that some resampling scheme must be added into the model, for the sake of synchronizing both processes.

7.3.3. Lossy compression of tree structures

Participant: Jean-Baptiste Durand.

Joint work with: Christophe Godin (Inria, Virtual Plants) and Romain Azais (Inria BIGS)

In a previous work [65], a method to compress tree structures and to quantify their degree of self-nestedness was developed. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG (Directed Acyclic Graph), equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the lossless compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree. A lossy version of the algorithm consists in coding the nearest self-nested tree embedded in the initial tree. Indeed, finding the nearest self-nested tree of a structure without more assumptions is conjectured to be an NP-complete or NP-hard problem. We improved this lossy compression method by computing a self-nested reduction of a tree that better approximates the initial tree. The algorithm has polynomial time complexity for trees with bounded outdegree. This approximation relies on an indel edit distance that allows (recursive) insertion and deletion of leaf vertices only. We showed in a conference paper accepted at DCC'2016 [55] with a simulated dataset that the error rate of this lossy compression method is always better than the loss based on the nearest embedded self-nestedness tree [65] while the compression rates are equivalent. This procedure is also a keystone in our new topological clustering algorithm for trees. In addition, we obtained new theoretical results on the combinatorics of self-nested structures. The redaction of an article is currently in progress.

7.4. Statistical models for Neuroscience

7.4.1. Comparison of stochastic and variational solutions to ASL fMRI data analysis

Participants: Florence Forbes, Aina Frau Pascual.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA Saclay.

Functional Arterial Spin Labeling (fASL) MRI can provide a quantitative measurement of changes of cerebral blood flow induced by stimulation or task performance. fASL data is commonly analysed using a general linear model (GLM) with regressors based on the canonical hemodynamic response function. In this work [37], we consider instead a joint detection-estimation (JDE) framework which has the advantage of allowing the extraction of both task-related perfusion and hemodynamic responses not restricted to canonical shapes. Previous JDE attempts for ASL have been based on computer intensive sampling (MCMC) methods. Our contribution is to provide a comparison with an alternative variational expectation-maximization (VEM) algorithm on synthetic and real data. Other investigations were related to the use of appropriate physiological information and priors [39], [38].

7.4.2. A differential evolution-based approach for fitting a nonlinear biophysical model to fMRI BOLD data

Participants: Florence Forbes, Pablo Mesejo.

Joint work with: Jan Warnking from Grenoble Institute of Neuroscience.

Physiological and biophysical models have been proposed to link neuronal activity to the Blood Oxygen Level-Dependent (BOLD) signal in functional MRI (fMRI). Those models rely on a set of parameter values that cannot always be extracted from the literature. In some applications, interesting insight into the brain physiology or physiopathology can be gained from an estimation of the model parameters from measured BOLD signals. This estimation is challenging because there are more than 10 potentially interesting parameters involved in nonlinear equations and whose interactions may result in identifiability issues. However, the availability of statistical prior knowledge about these parameters can greatly simplify the estimation task. In this work we focus on the extended Balloon model and propose the estimation of 15 parameters using two stochastic approaches: an Evolutionary Computation global search method called Differential Evolution (DE)

and a Markov Chain Monte Carlo version of DE. To combine both the ability to escape local optima and to incorporate prior knowledge, we derive the target function from Bayesian modeling. The general behavior of these algorithms is analyzed and compared with the *de facto* standard Expectation Maximization Gauss-Newton (EM/GN) approach, providing very promising results on challenging real and synthetic fMRI data sets involving rats with epileptic activity. These stochastic optimizers provided a better performance than EM/GN in terms of distance to the ground truth in 4 out of 6 synthetic data sets and a better signal fitting in 12 out of 12 real data sets. Non-parametric statistical tests showed the existence of statistically significant differences between the real data results obtained by DE and EM/GN. Finally, the estimates obtained from DE for these parameters seem both more realistic and more stable or at least as stable across sessions as the estimates from EM/GN. This work will appear in [29]. A preliminary version has also been accepted at the conference MICCAI 2015 [40].

7.4.3. Multi-subject joint parcellation detection estimation in functional MRI

Participant: Florence Forbes.

Joint work with: Lotfi Chaari, Mohanad Albughdadi, Jean-Yves Tourneret from IRIT-ENSEEIH in Toulouse and Philippe Ciuciu from Neurospin, CEA Saclay.

fMRI experiments are usually conducted over a population of interest for investigating brain activity across different regions, stimuli and subjects. Multi-subject analysis usually proceeds in two steps: an intra-subject analysis is performed sequentially on each individual and then a group-level analysis is carried out to report significant results at the population level. This work considers an existing Joint Parcellation Detection Estimation (JPDE) model which performs joint hemodynamic parcellation, brain dynamics estimation and evoked activity detection. The hierarchy of the JPDE model is extended for multi-subject analysis in order to perform group-level parcellation. Then, the corresponding underlying dynamics is estimated in each parcel while the detection and estimation steps are iterated over each individual. Validation on synthetic and real fMRI data shows its robustness in inferring group-level parcellation and the corresponding hemodynamic profiles. This work has been accepted at ISBI 2016.

7.4.4. Tumor classification and prediction using robust multivariate clustering of multiparametric MRI

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Emmanuel Barbier and Benjamin Lemasson from Grenoble Institute of Neuroscience.

Advanced statistical clustering approaches are promising tools to better exploit the wealth of MRI information especially on large cohorts and multi-center studies. In neuro-oncology, the use of multiparametric MRI may better characterize brain tumor heterogeneity. To fully exploit multiparametric MRI (e.g. tumor classification), appropriate analysis methods are yet to be developed. They offer improved data quality control by allowing automatic outlier detection and improved analysis by identifying discriminative tumor signatures with measurable predictive power. In this work, we show on small animals data that advanced statistical learning approaches can help 1) in organizing existing data by detecting and excluding outliers and 2) in building a dictionary of tumor fingerprints from a clustering analysis of their microvascular features. The work also now includes the integration in a joint statistical model of both automatic ROI delineation and clustering for whole brain data analysis. A preliminary version of this work has been accepted to the ISMRM 2015 conference and in the SFMRMB 2015 conference [41].

7.4.5. Functional specifications of a brain segmentation software

Participants: Florence Forbes, Priscillia Previtro.

Joint work with: Michel Dojat from Grenoble Institute of Neuroscience and Senan Doyle from Pixyl.

The goal of P. Previtro's internship was to help with a number of software engineering tasks and communications actions around the P-Locus software and the Pixyl start-up. The internship resulted in particular in a new web site for Pixyl.

MODAL Project-Team

7. New Results

7.1. Functional data analysis applied to hydrological or environmental data

Participant: Sophie Dabo.

The new results concern particularly functional data analysis applied to hydrological or environmental data. First in a recent paper ([16]), two statistical techniques from the theory of functional data classification are adapted and applied for the analysis of flood hydrographs. Functional classification directly employs all data of a discharge time series and thus contains all available information on shape, peak, and timing. This potentially allows a better understanding and treatment of floods as well as other hydrological phenomena.

7.2. New functional regression model when data are auto-correlated

Participant: Sophie Dabo.

We develop a new functional regression model when data are auto-correlated, in collaboration with Serge Guillas (University of College London) and Camille Ternynck (University of Lille 2). This work will appear in *Journal of Multivariate Analysis*. (Dabo-Niang, S, Guillas, S et Ternynck, C. (2016). More efficient kernel functional spatial regression estimation with autocorrelated errors. *Journal of Multivariate Analysis*). In this work we introduce a new procedure for the estimation in the nonlinear functional regression model where the explanatory variable takes values in an abstract function space and the residual process is autocorrelated. The procedure consists in a pre-whitening transformation of the dependent variable based on the estimated autocorrelation. We establish both consistency and asymptotic normality of the regression function estimate. For kernel methods encountered in the literature, the correlation structure is commonly ignored (the so-called “working independence estimator”); we show here that there is a strong benefit in taking into account the autocorrelation in the error process. We also find that the improvement in efficiency can be large in our functional setting, up to 25% in the presence of high autocorrelation levels. Concerning spatial data, we develop a new spatial prediction method that takes into account the spatial dependence. This work will appear in *Journal of Nonparametric Statistics* (Dabo-Niang, Ternynck, C., Yao, A.-F. (2016). Nonparametric prediction in the multivariate spatial context. *Journal of Nonparametric Statistics*)

7.3. Differential gene expression analysis

Participant: Guillemette Marot.

The use of empirical Bayesian techniques implemented in the Bioconductor package `limma` has enabled to better understand Waldenström’s macroglobulinemia. Gene Set enrichment analysis was also performed after differential analysis. The new findings in Biology have been published in [21].

7.4. Evolutionary clustering for categorical data

Participant: Julien Jacques.

This is a joint work with Md Abul Hasnat, Julien Velcin and Stephane Bonnevey (Univ. de Lyon).

An evolutionary clustering algorithm for categorical data has been developed, based on parametric links between multinomial mixture models. This model has been used to study the evolution of opinions in Twitter data. A Preprint of this work is available [54].

7.5. Clustering categorical functional data: Application to medical discharge letters

Participants: Cristian Preda, Cristina Preda, Vincent Vandewalle.

Categorical functional data represented by paths of a stochastic jump process are considered for clustering. For paths of the same length, the extension of the multiple correspondence analysis allows the use of well-known methods for clustering finite dimensional data. When the paths are of different lengths, the analysis is more complex. In this case, for Markov models we have proposed an EM algorithm to estimate a mixture of Markov processes. This work has been presented in a conference [34].

7.6. Degeneracy in Gaussian Mixtures with missing data

Participants: Christophe Biernacki, Vincent Vandewalle.

The missing data problem is well-known for statisticians but its frequency increases with the growing size of modern datasets. In Gaussian model-based clustering, the EM algorithm easily takes into account such data by dealing with two kinds of latent levels: the components and the variables. However, the quite familiar degeneracy problem in Gaussian mixtures is aggravated during the EM runs. Indeed, numerical experiments clearly reveal that degeneracy is quite slow and also more frequent than with complete data. In practice, such situations are difficult to detect efficiently. Consequently, degenerated solutions may be confused with valuable solutions and, in addition, computing time may be wasted through wrong runs. A simple condition on the latent partition to avoid degeneracy has been exhibited, and a constrained version of the Stochastic EM (SEM) algorithm satisfying this condition has been proposed. This work has been presented in a conference [33].

7.7. Model for conditionally correlated categorical data

Participants: Christophe Biernacki, Vincent Vandewalle, Matthieu Marbac-Lourdelle.

An extension of the latent class model is proposed for clustering categorical data by relaxing the classical class conditional independence assumption of variables. In this model (called CCM for Conditional Correlated Model), variables are grouped into inter-independent and intra-dependent blocks in order to consider the main intra-class correlations. The dependence between variables grouped into the same block is taken into account by mixing two extreme distributions, which are respectively the independence and the maximum dependence ones. In the conditionally correlated data case, this approach is expected to reduce biases involved by the latent class model and to produce a meaningful model with few additional parameters. The parameters estimation by maximum likelihood is performed by an EM algorithm while a MCMC algorithm avoiding combinatorial problems involved by the block structure search is used for model selection. Applications on sociological and biological data sets bring out the proposed model interest. These results strengthen the idea that the proposed model is meaningful and that biases induced by the conditional independence assumption of the latent class model are reduced. This work is published [20]. Furthermore, an R package (Clustericat) is available on Rforge (see <https://github.com/rforge/clustericat>).

7.8. Model-based clustering for multivariate partial ordinal data

Participants: Christophe Biernacki, Julien Jacques.

We design the first univariate probability distribution for ordinal data which strictly respects the ordinal nature of data. More precisely, it relies only on order comparisons between modalities, the proposed distribution being obtained by modeling the data generating process which is assumed, from optimality arguments, to be a stochastic binary search algorithm in a sorted table. The resulting distribution is natively governed by two meaningful parameters (position and precision) and has very appealing properties: decrease around the mode, shape tuning from uniformity to a Dirac, identifiability. Moreover, it is easily estimated by an EM algorithm since the path in the stochastic binary search algorithm is missing. Using then the classical latent class assumption, the previous univariate ordinal model is straightforwardly extended to model-based clustering for multivariate ordinal data. Again, parameters of this mixture model are estimated by an EM algorithm. Both simulated and real data sets illustrate the great potential of this model by its ability to parsimoniously identify particularly relevant clusters which were unsuspected by some traditional competitors. This work is now published in an international journal [12] and is also currently available in the MixtComp software at <https://modal-research.lille.inria.fr/BigStat/>

7.9. Semi-Linear Auto-Associative Model

Participant: Serge Iovleff.

We design a new model for data analysis which is a generalization of the probabilistic PCA. The interpretation properties of the PCA are preserved while presence of non-linear repartitions in data can be detected and adjusted using B-spline regression. This model has been published in [18].

MOKAPLAN Project-Team

7. New Results

7.1. Numerical methods for JKO Gradient Flows

J-D. Benamou, G. Carlier, M. Laborde, G. Peyré, B. Schmitzer, V. Duval

Taking advantage of the Benamou-Brenier dynamic formulation of optimal transport, we propose in [28], a convex formulation for each step of the JKO scheme for Wasserstein gradient flows which can be attacked by an augmented Lagrangian method which we call the ALG2-JKO scheme. We test the algorithm in particular on the porous medium equation. We also consider a semi implicit variant which enables us to treat nonlocal interactions as well as systems of interacting species. Regarding systems, we can also use the ALG2-JKO scheme for the simulation of crowd motion models with several species.

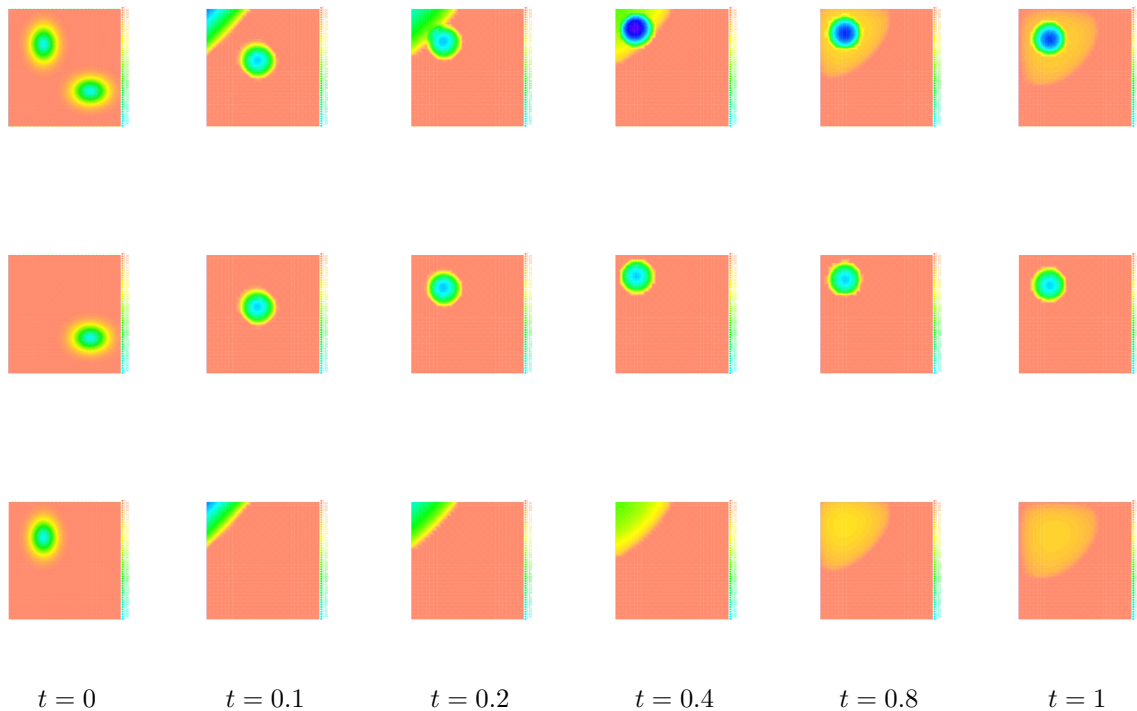


Figure 12. Evolution of two species where the first one is attracted by the other and the second one is repelled by the first one. Top row: display of $\rho_1 + \rho_2$. Middle row: display of ρ_1 . Bottom row: display of ρ_2 .

We have also investigated the entropy-regularization of the Wasserstein metric to compute gradient flows [19], [34]. This entropic regularization trades the usual Wasserstein fidelity term for a Kullback-Leibler divergence term. Adapting first-order proximal methods to this framework, we have developed numerical schemes which dramatically reduce the computational load needed to simulate the evolution of a mass density through a JKO flow. By construction, the entropy regularization yields an additional diffusion effects to the evolution, but we have proved that a careful choice of the regularization parameter with respect to the timestep yields the convergence of the scheme towards the solutions of the continuous PDE.

A novel Lagrangian method using a discretization of the Monge-Ampère operator for JKO has been developed in [13]. Not only convergence of the scheme has been established but also one advantage of this method is that it makes it possible to use a Newton's method .

7.2. Density Functional Theory

J-D. Benamou Luca Nenna, G. Carlier

In [41] is presented the state of art and recent developments of the optimal transportation theory with many marginals for a class of repulsive cost functions. We introduce some aspects of the Density Functional Theory (DFT) from a mathematical viewpoint, and revisit the theory of optimal transport from its perspective. Moreover, in the last three sections, we describe some recent and new theoretical and numerical results obtained for the Coulomb cost, the repulsive harmonic cost and the determinant.

In [29] we present a numerical method, based on iterative Bregman projections, to solve the optimal transport problem with Coulomb cost. This is related to the strong interaction limit of Density Functional Theory. The first idea is to introduce an entropic regularization of the Kantorovich formulation of the Optimal Transport problem. The regularized problem then corresponds to the projection of a vector on the intersection of the constraints with respect to the Kullback-Leibler distance. Iterative Bregman projections on each marginal constraint are explicit which enables us to approximate the optimal transport plan. We validate the numerical method against analytical test cases.

7.3. Stability for inverse problems with sparsity prior

G. Peyré, V. Duval, Q. Denoyelle, C. Poon

In [42], we have analyzed the recovery performance of two popular finite dimensional approximations of the sparse spikes deconvolution problem over Radon measures, namely the LASSO, and the Continuous Basis-Pursuit. The LASSO is the de-facto standard for the sparse regularization of inverse problems in imaging. It performs a nearest neighbor interpolation of the spikes locations on the sampling grid. The C-BP method, introduced by Ekanadham, Tranchina and Simoncelli, uses a linear interpolation of the locations to perform a better approximation of the infinite-dimensional optimization problem, for positive measures. We have proved that, in the small noise regime, both methods estimate twice the number of original spikes, and we have provided an explicit formula which allows to predict the locations and amplitudes of the spurious spikes. All those properties are in fact connected to an intrinsic property of the signal: the source condition [16], [24].

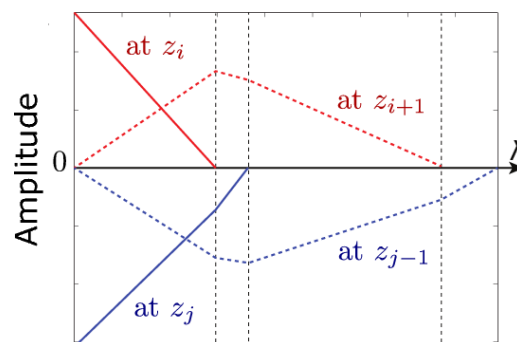


Figure 13. The solution path of the discrete LASSO (as a function of λ) for some discrete measure m_0 (the noise w is set to zero). This shows the amplitudes of the coefficients at $z_i = ih$, resp. $z_j = jh$, (continuous line) and at the next, resp. previous, point of the grid (dashed line) as λ varies.

Those effects are typically due to the use of a discrete grid in the reconstruction process. Several authors have recently proposed algorithms to tackle the problem directly in a continuous setting [75], [92]. As we have shown in [16], the method fails when the distance between spikes with opposite signs are below a certain threshold. However, when all the spikes have the same sign, the LASSO on a continuous domain works for arbitrarily close spikes, being all the more sensitive to noise. In [40], we have given a detailed analysis of the noise sensitivity of the method: if t denotes the minimum separation of the input measure (the minimum distance between two spikes), w refers to the noise and λ is the regularization parameter, when $\|w\|_{L^2}/\lambda$, $\|w\|_{L^2}/t^{2N-1}$ and λ/t^{2N-1} are small enough (where N is the number of spikes), there exists a unique solution to the BLASSO program with exactly the same number of spikes as the original measure. We show that the amplitudes and positions of the spikes of the solution both converge toward those of the input measure when the noise and the regularization parameter drops to zero faster than t^{2N-1} .

7.4. Generalized Solution of Euler

Minimal geodesics along volume preserving maps, through semi-discrete optimal transport

Q. Mérigot and J.-M. Mirebeau introduced a numerical method for extracting minimal geodesics along the group of volume preserving maps, equipped with the L^2 metric, which as observed by Arnold solve Euler's equations of inviscid incompressible fluids. The method relies on the generalized polar decomposition of Brenier, numerically implemented through semi-discrete optimal transport. It is robust enough to extract non-classical, multi-valued solutions of Euler's equations, for which the dimension of the support of the flow is higher than the dimension of the domain, a striking and unavoidable consequence of this model. Our convergence results encompass this generalized model, and our numerical experiments illustrate it for the first time in two space dimensions (see Figure 14).

7.5. Principal Agent Problem

J.-D. Benamou, Xavier Dupuis, G. Carlier An alternated projection numerical scheme for the more general c -concavity constraint using Dykstra's algorithm has been recently developed in [33] but being able to handle realistic principal-agent problems remains a challenging issue. Investigating the structure of equilibria in matching problems with non-transferable utilities is also one of our objectives, together with numerical methods in the spirit of the IPFP algorithm.

A semi-discrete approach to the PA problem is investigated. The range of products is discrete and leads to a non convex problem. Non-linear optimization methods are tested. See <https://mathmarx.paris.inria.fr:8080>.

7.6. Unbalanced Optimal Transport

G. Carlier, F.-X. Vialard, B. Schmitzer, L. Chizat Classical optimal transport theory and algorithms assume that the input measures are normalized, i.e. that their total mass is 1. This is an important limitation for many problems in imaging sciences and machine learning, where input data are typically not normalized, and where one should enables local creation or destruction of mass. Handling such "unbalanced" transportation problem is also relevant for applications in biological modeling, for instance to take into account cellular growth through optimal transport gradient flows.

Recently, several researchers of MOKAPLAN made important progress on this problem, by deriving a general framework extending optimal transport to this "unbalanced" setting. In [38] we derived a dynamic optimal transport formulation that enables a source term in the initial formulation of Benamou and Brenier [55]. We proved that it defines a distance on positive measures, enjoy many important properties (dual formulation) and can be computed using fast first order convex optimization methods. We then provided in [39] an even larger class of "unbalanced" optimal transport optimization problems, that are obtained via a static formulation, and show that one can recovers the dynamic formulation in some specific cases. Similar models were derived independently and at the same time by two other international teams [143], [137], which shows the timeliness of our research. We believe these new theoretical and numerical findings will have a strong impact on the developpement of optimal transport methods in imaging sciences and machine learning.

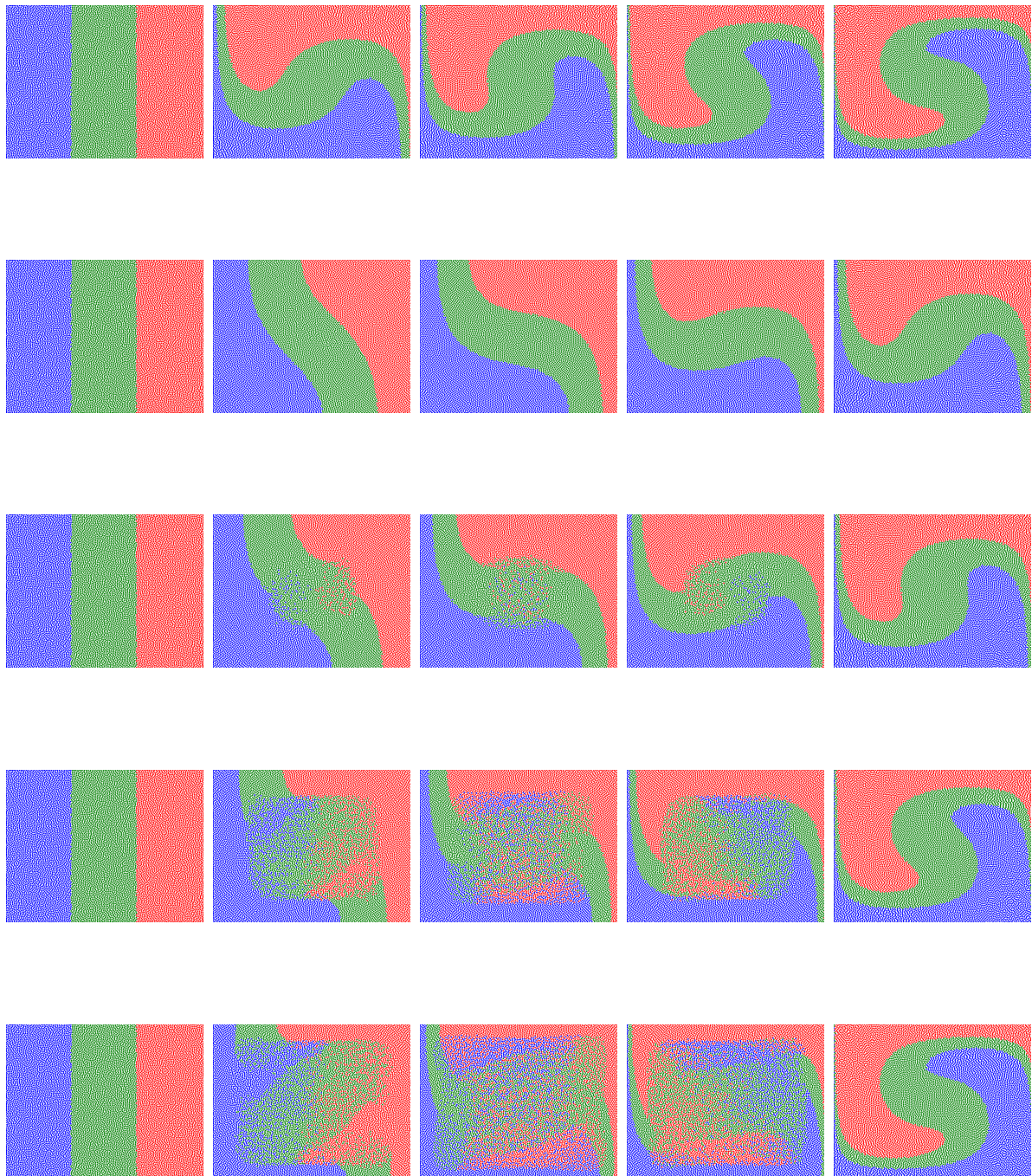


Figure 14. (First row) Beltrami flow in the unit square at various timesteps, a classical solution to Euler's equation. The color of the particles depend on their initial position. (Second to fifth row) Generalized fluid flows that are reconstructed by our algorithm, using boundary conditions displayed in the first and last column. When $t_{\max} < 1$ we recover the classical flow, while for $t_{\max} \geq 1$ the solution is not classical any more and includes some mixing.

NACHOS Project-Team

6. New Results

6.1. Electromagnetic wave propagation

6.1.1. Numerical study of the non-linear Maxwell equations for Kerr media

Participants: Loula Fezoui, Stéphane Lanteri.

The system of Maxwell equations describes the evolution of the interaction of an electromagnetic field with a propagation medium. The different properties of the medium, such as isotropy, homogeneity, linearity, among others, are introduced through *constitutive laws* linking fields and inductions. In the present study, we focus on non-linear effects and address non-linear Kerr materials specifically. In this model, any dielectric may become non-linear provided the electric field in the material is strong enough. As a first step, we considered the one-dimensional case and study the numerical solution of the non-linear Maxwell equations thanks to DG methods. In particular, we make use of an upwind scheme and limitation techniques because they have a proven ability to capture shocks and other kinds of singularities in the fluid dynamics framework. The numerical results obtained in this preliminary study gave us confidence towards extending them to higher spatial dimensions. As a matter of fact, we recently started to work on the three-dimensional case and have initiated the development of a parallel simulation software based on our past contributions on DGTD methods for the case of linear propagation media.

6.1.2. High order geometry conforming DGTD method for nanophotonics

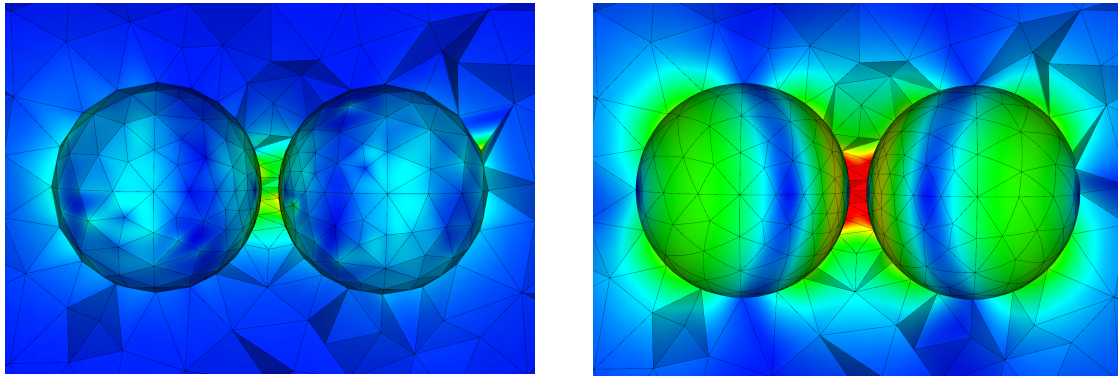
Participants: Stéphane Lanteri, Claire Scheid, Jonathan Viquerat.

Usually, unstructured mesh based methods rely on tessellations composed of straight-edged elements mapped linearly from a reference element, on domains which physical boundaries are indifferently straight or curved. Such meshes represent a serious hindrance for high order finite element (FE) methods since they limit the accuracy to second order in the spatial discretization. Thus, exploiting an enhanced representation of physical geometries is in agreement with the natural procedure of high order FE methods, such as the DG method. There are several ways to account for curved geometries. One could choose to incorporate the knowledge coming from CAD in the method to design the geometry and the approximation. These methods are called *isogeometric*, and have received a lot of attention recently. This naturally implies to have access to CAD models of the geometry. On the other hand, *isoparametric* usually rely on a polynomial approximation of both the boundary and the solution. This can be added fairly easily on top of existing implementations. In the present study we focus on the latter type of method, since our goal is first to envisage the benefit of curvilinear meshes for light/matter interaction with nanoscale structures.

6.1.3. Local approximation order DGTD method for nanophotonics

Participants: Stéphane Lanteri, Jonathan Viquerat.

High order DGTD methods for the numerical modeling of light/matter interactions on the nanoscale often assume a uniform distribution of the polynomial order to the cells of the underlying mesh. However, in the case of a mesh showing large variations in cell size, the time step imposed by the smallest cells can be a serious hindrance when trying to exploit high approximation orders. Indeed, a potentially large part of the CPU time is spent in the update of the physical field inside small cells where high polynomial orders might not be essential, while they are necessary in the larger cells. In this study, we consider the possibility of using a non-uniform distribution of the polynomial order in the framework of a global time step DGTD method. By imposing low orders in small cells and high orders in large cells, it is possible to significantly alleviate both the global number of degrees of freedom and the time step restriction with a minimal impact on the method accuracy. Strategies exploiting locally adaptive (LA) formulations usually combine both h - and p -adaptivity (where h denotes the discretization parameter in space and p the degree of the interpolation of the field components) in



Mesh with affine elements

Mesh with curvilinear elements

Figure 4. Near-field visualization of the amplitude of the electric field Fourier transformed for a gold nanosphere dimer. Surface-to-surface distance is set to 4 nm. Calculations are based on a DGTD- \mathbb{P}_4 method.

order to concentrate the computational effort in the areas of high field variations. Here, the adopted point of view is quite different: starting from a given mesh and a uniform distribution of the polynomial order k , the LA strategy exploits all the polynomial orders p with $p \leq k$ to obtain a solution of similar accuracy with a reduced computational cost.

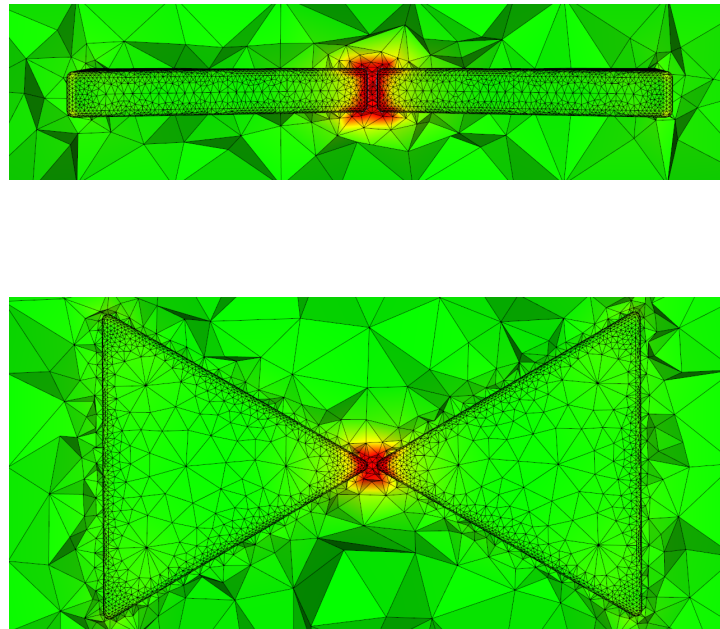


Figure 5. Near-field plasmonic interaction with a bowtie nanoantennas: contour line the amplitude of the discrete Fourier transform of the electric field.

6.1.4. Numerical treatment of non-local dispersion for nanoplasmonics

Participants: Stéphane Lanteri, Claire Scheid, Colin Vo Cong Tri.

When metallic nanostructures have sub-wavelength sizes and the illuminating frequencies are in the regime of metal's plasma frequency, electron interaction with the exciting fields have to be taken into account. Due to these interactions, plasmonic surface waves can be excited and cause extreme local field enhancements (surface plasmon polariton electromagnetic waves). Exploiting such field enhancements in applications of interest requires a detailed knowledge about the occurring fields which can generally not be obtained analytically. For the numerical modeling of light/matter interaction on the nanoscale, the choice of an appropriate model is a crucial point. Approaches that are adopted in a first instance are based on local (no interaction between electrons) dispersion models e.g. Drude or Drude-Lorentz. From the mathematical point of view, these models lead to an additional ordinary differential equation in time that is coupled to Maxwell's equations. When it comes to very small structures in a regime of 2 nm to 25 nm, non-local effects due to electron collisions have to be taken into account. Non-locality leads to additional, in general non-linear, partial differential equations and is significantly more difficult to treat, though. In this work, we study a DGTD method able to solve the system of Maxwell equations coupled to a linearized non-local dispersion model relevant to nanoplasmonics. While the method is presented in the general 3d case, in this preliminary study, numerical results are given for 2d simulation settings.

6.1.5. Corner effects in nanoplasmonics

Participants: Camille Carvalho [ENSTA, POEMS project-team], Patrick Ciarlet [ENSTA, POEMS project-team], Claire Scheid.

The starting point of this ongoing work is the theoretical and numerical study of nanoplasmonic structures with corners. This is the central subject of the PhD thesis of Camille Carvalho. In the latter, the focus is made on a lossless Drude dispersion model with a frequency domain approach. Several well posedness problems arise due to the presence of corners and are addressed in the PhD thesis. A time domain approach in this context can also be investigated with the techniques developed in our project-team. Even if both approaches (time domain and frequency domain) represent similar physical phenomena, problems that arise are different and they appear as complementary; it is thus worth bridging the gap between the two frameworks. We propose to perform a thorough comparison in the case of structures with corners. Several extensions to other models are also envisaged, especially concerning the non local dispersion model.

6.1.6. DGTD method for nanoplasmonics based on generalized dispersion model

Participants: Stéphane Lanteri, Claire Scheid, Jonathan Viquerat.

In this work, we are concerned with the numerical modelling of the propagation of electromagnetic waves in dispersive materials for nanophotonics applications. We focus on a generalized model that allows for the description of a wide range of dispersive media. The underlying differential equations are recast into a generic form and we establish an existence and uniqueness result. We then turn to the numerical treatment and propose an appropriate DGTD framework. We obtain the semi-discrete convergence and prove the stability (and to a large extent, the convergence) in the fully discrete case when time integration is achieved with a 4 steps low storage Runge-Kutta scheme, *via* a technique relying on energy principles. Finally, we validate our results through the numerical simulation of two nanophotonics test cases.

6.1.7. Travelling waves for the non-linear Schrödinger equation in 2d

Participants: David Chiron [J.A. Dieudonné Laboratory, University of Nice-Sophia Antipolis], Claire Scheid.

We are interested in the numerical study of the two-dimensional travelling waves of the non-linear Schrödinger equation for a general non-linearity and with nonzero condition at infinity that can appear in optics. The equation has a variational structure that we propose to exploit to design a numerical method. We characterize the saddle points of the action as minimizers of another functional, allowing us to use a gradient flow. Combining this approach with a continuation method in the speed of the wave, we obtain the numerical

solution for the full range of velocities. We plot the energy-momentum diagrams for different type of non-linearities. Through various examples, we show that even though the non-linearity has the same behaviour as the well-known Gross-Pitaevskii (GP) non-linearity, the qualitative properties of the travelling waves may be extremely different. For instance, we observe cusps, a modified Kadomtsev-Petviashvili I (KP-I) asymptotic in the transonic limit (as the speed of the wave approaches the speed of sound), various multiplicity results.

6.1.8. Multiscale DG methods for the time-domain Maxwell equations

Participants: Stéphane Lanteri, Raphaël Léger, Diego Paredes Concha [Instituto de Matemáticas, Universidad Católica de Valparaíso, Chile], Claire Scheid, Frédéric Valentin [LNCC, Petropolis, Brazil].

Although the DGT method has already been successfully applied to complex electromagnetic wave propagation problems, its accuracy may seriously deteriorate on coarse meshes when the solution presents multiscale or high contrast features. In other physical contexts, such an issue has led to the concept of multiscale basis functions as a way to overcome such a drawback and allow numerical methods to be accurate on coarse meshes. The present work, which has been initiated in the context of the visit of Frédéric Valentin in the team, is concerned with the study of a particular family of multiscale methods, named Multiscale Hybrid-Mixed (MHM) methods. Initially proposed for fluid flow problems, MHM methods are a consequence of a hybridization procedure which characterizes the unknowns as a direct sum of a coarse (global) solution and the solutions to (local) problems with Neumann boundary conditions driven by the purposely introduced hybrid (dual) variable. As a result, the MHM method becomes a strategy that naturally incorporates multiple scales while providing solutions with high order accuracy for the primal and dual variables. The completely independent local problems are embedded in the upscaling procedure, and computational approximations may be naturally obtained in a parallel computing environment. In this study, a family of MHM methods is proposed for the solution of the time-domain Maxwell equations where the local problems are discretized either with a continuous FE method or a DG method (that can be viewed as a multiscale DGT method). Preliminary results have been obtained in the 2d case for models problems.

6.1.9. HDG methods for the time-domain Maxwell equations

Participants: Alexandra Christophe-Argenvillier, Stéphane Descombes, Stéphane Lanteri.

This study is concerned with the development of accurate and efficient solution strategies for the system of 3d time-domain Maxwell equations coupled to local dispersion models (e.g. Debye, Drude or Drude-Lorentz models) in the presence of locally refined meshes. Such meshes impose a constraint on the allowable time step for explicit time integration schemes that can be very restrictive for the simulation of 3d problems. We consider here the possibility of using an unconditionally stable implicit time or a locally implicit time integration scheme combined to a HDG discretization method. In doing so, we extend our former study which was dealing with the 2d time-domain Maxwell equations for non-dispersive media.

6.1.10. HDG methods for the frequency-domain Maxwell equations

Participants: Thomas Frachon, Stéphane Lanteri, Liang Li [UESTC, Chengdu, China], Ludovic Moya, Ronan Perrussel [Laplace Laboratory, Toulouse].

In the context of the ANR TECSER project, we continue our efforts towards the development of scalable high order HDG methods for the solution of the system of 3d frequency-domain Maxwell equations. We aim at fully exploiting the flexibility of the HDG discretization framework with regards to the adaptation of the interpolation order (p -adaptivity) and the mesh (h -adaptivity). In particular, we study the formulation of HDG methods on a locally refined non-conforming tetrahedral mesh and on a non-conforming hybrid cubic/tetrahedral mesh. We also investigate the coupling between the HDG formulation and a BEM (Boundary Element Method) discretization of an integral representation of the electromagnetic field in the case of propagation problems theoretically defined in unbounded domains.

6.2. Elastodynamic wave propagation

6.2.1. Seismic wave interaction with viscoelastic media

Participants: Nathalie Glinsky, Stéphane Lanteri, Fabien Peyrusse [Department of Mathematics, Purdue University].

This work is concerned with the development of high order DGTD methods formulated on unstructured simplicial meshes for the numerical solution of the system of time-domain elastodynamic equations. These methods share some ingredients of the DGTD methods developed by the team for the time-domain Maxwell equations among which, the use of nodal polynomial (Lagrange type) basis functions, a second order leap-frog time integration scheme and a centered scheme for the evaluation of the numerical flux at the interface between neighboring elements. A recent novel contribution is the numerical treatment of viscoelastic attenuation. For this, the velocity-stress first order hyperbolic system is completed by additional equations for the anelastic functions including the strain history of the material. These additional equations result from the rheological model of the generalized Maxwell body and permit the incorporation of realistic attenuation properties of viscoelastic material accounting for the behaviour of elastic solids and viscous fluids. In practice, we need solving $3L$ additional equations in 2d (and $6L$ in 3d), where L is the number of relaxation mechanisms of the generalized Maxwell body. This method has been implemented in 2d and 3d.

6.2.2. DG method for arbitrary heterogeneous media

Participants: Nathalie Glinsky, Diego Mercerat [CETE Méditerranée].

We have recently devised an extension of the DGTD method for elastic wave propagation in arbitrary heterogeneous media. In realistic geological media (sedimentary basins for example), one has to include strong variations in the material properties. Then, the classical hypothesis that these properties are constant within each element of the mesh can be a severe limitation of the method, since we need to discretize the medium with very fine meshes resulting in very small time steps. For these reasons, we propose an improvement of the DGTD method allowing non-constant material properties within the mesh elements. A change of variables on the stress components allows writing the elastodynamic system in a pseudo-conservative form. Then, the introduction of non-constant material properties inside an element is simply treated by the calculation, via convenient quadrature formulae, of a modified local mass matrix depending on these properties. This new extension has been validated for a smoothly varying medium or a strong jump between two media, which can be accurately approximated by the method, independently of the mesh.

6.2.3. HDG method for the frequency-domain elastodynamic equations

Participants: Hélène Barucq [MAGIQUE-3D project-team, Inria Bordeaux - Sud-Ouest], Marie Bonnasse, Julien Diaz [MAGIQUE-3D project-team, Inria Bordeaux - Sud-Ouest], Stéphane Lanteri.

One of the most used seismic imaging methods is the full waveform inversion (FWI) method which is an iterative procedure whose algorithm is the following. Starting from an initial velocity model, (1) compute the solution of the wave equation for the N sources of the seismic acquisition campaign, (2) evaluate, for each source, a residual defined as the difference between the wavefields recorded at receivers on the top of the subsurface during the acquisition campaign and the numerical wavefields, (3) compute the solution of the wave equation using the residuals as sources, and (4) update the velocity model by cross correlation of images produced at steps (1) and (3). Steps (1)-(4) are repeated until convergence of the velocity model is achieved. We then have to solve $2N$ wave equations at each iteration. The number of sources, N , is usually large (about 1000) and the efficiency of the inverse solver is thus directly related to the efficiency of the numerical method used to solve the wave equation. Seismic imaging can be performed in the time-domain or in the frequency-domain regime. In this work which is conducted in the framework of the Depth Imaging Partnership (DIP) between Inria and TOTAL, we adopt the second setting. The main difficulty with frequency-domain inversion lies in the solution of large sparse linear systems which is a challenging task for realistic 3d elastic media, even with the progress of high performance computing. In this context, we study novel high order HDG methods formulated on unstructured meshes for the solution of the frequency-domain elastodynamic equations. Instead of solving a linear system involving the degrees of freedom of all volumic cells of the mesh, the principle of a HDG formulation is to introduce a new unknown in the form of Lagrange multiplier representing the trace of the numerical solution on each face of the mesh. As a result, a HDG formulation yields a global linear system in terms of the new (surfacic) unknown while the volumic solution is recovered thanks to a local computation on each element.

6.2.4. Multiscale DG methods for the time-domain elastodynamic equations

Participants: Marie-Hélène Lallemand, Frédéric Valentin [LNCC, Petropolis, Brazil].

In the context of the visit of Frédéric Valentin in the team, we have initiated a study aiming at the design of novel multiscale methods for the solution of the time-domain elastodynamic equations, in the spirit of MHM (Multiscale Hybrid-Mixed) methods previously proposed for fluid flow problems. Motivation in that direction naturally came when dealing with non homogeneous anisotropic elastic media as those encountered in geodynamics related applications, since multiple scales are naturally present when high contrast elasticity parameters define the propagation medium. Instead of solving the usual system expressed in terms of displacement or displacement velocity, and stress tensor variables, a hybrid mixed-form is derived in which an additional variable, the Lagrange multiplier, is sought as representing the (opposite) of the surface tension defined at each face of the elements of a given discretization mesh. We consider the velocity/stress formulation of the elastodynamic equations, and study a MHM method defined for a heterogeneous medium where each elastic material is considered as isotropic to begin with. If the source term (the applied given force on the medium) is time independent, and if we are given an arbitrarily coarse conforming mesh (triangulation in 2d, tetrahedrization in 3d), the proposed MHM method consists in first solving a series of fully decoupled (therefore parallelizable) local (element-wise) problems defining parts of the full solution variables which are directly related to the source term, followed by the solution of a global (coarse) problem, which yields the degrees of freedom of both the Lagrange multiplier dependent part of the full solution variables and the Lagrange multiplier itself. Finally, the updating of the full solution variables is obtained by adding each splitted solution variables, before going on the next time step of a leap-frog time integration scheme. Theoretical analysis and implementation of this MHM method where the local problems are discretized with a DG method, are underway.

6.3. High performance numerical computing

6.3.1. Porting a DGTD solver for bioelectromagnetics to the DEEP-ER architecture

Participants: Alejandro Duran [Barcelona Supercomputing Center, Spain], Stéphane Lanteri, Raphaël Léger, Damian A. Mallón [Juelich Supercomputing Center, Germany].

We are concerned here with the porting of a Discontinuous Galerkin Time-Domain solver for computational bioelectromagnetics to the novel heterogeneous architecture proposed in the DEEP-ER european project on exascale computing. This architecture is based on a Cluster/Booster division concept (see Fig. 6). The Booster nodes are based on the Intel Many Integrated Core (MIC) architecture. Therefore, one objective of our efforts is the algorithmic adaptation of the DG kernels in order to leverage the vectorizing capabilities of the MIC processor. The other activities that are undertaken in the context of our contribution to this project aim at exploiting the software environments and tools proposed by DEEP-ER partners for implementing resiliency strategies and high performance I/O operations. In particular, the Cluster nodes are used for running some parts of the pre- and post-processing phases of the DGTD solver which do not lend themselves well to multithreading, as well as I/O intensive routines. One possibility to achieve this is to consider a model in which these less scalable and I/O phases are reverse-offloaded from Booster processes to Cluster processes in a one-to-one mapping. This is achieved by exploiting the OmpSs offload functionality, developed at Barcelona Supercomputing Center for the DEEP-ER platform. In future work, the OmpSs framework will also be leveraged to expose task-based parallelism and exploit task-based resilience.

6.3.2. Hybrid MIMD/SIMD high order DGTD solver for nanophotonics

Participants: Tristan Cabel, Gabriel Hautreux [CINES, Montpellier], Stéphane Lanteri, Raphaël Léger, Claire Scheid, Jonathan Viquerat.

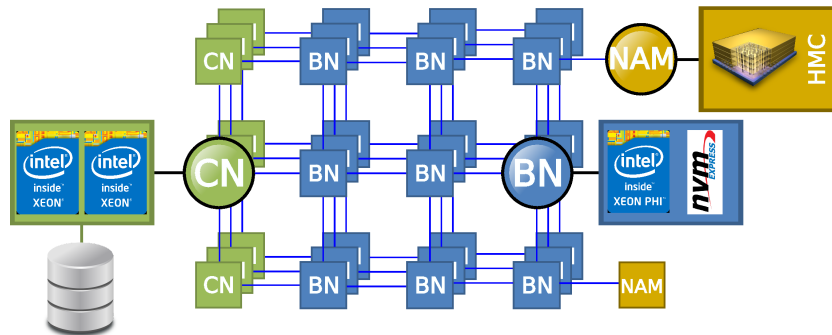


Figure 6. DEEP-ER hardware architecture sketch.

This work is concerned with the development of a scalable high order finite element type solver for the numerical modeling of light interaction with nanometer scale structures. From the mathematical modeling point of view, one has to deal with the differential system of Maxwell equations in the time domain, coupled to an appropriate differential model of the behavior of the underlying material (which can be a dielectric and/or a metal) at optical frequencies. For the numerical solution of the resulting system of differential equations, we adopt the high order DGTD (Discontinuous Galerkin Time-Domain) solver described in [21]. A hybrid MIMD/SIMD parallelization of this DGTD solver has been developed by combining the MPI and OpenMP parallel programming models. The performances of the resulting parallel DGTD solver have been assessed on the Curie system of the PRACE research infrastructure. For that purpose, we selected a use case typical of optical guiding applications. A Y-shaped waveguide is considered which consists in nanosphere embedded in vacuum (see Fig. 7). The constructed tetrahedral mesh consists of 520,704 vertices and 2,988,103 elements. The high order discontinuous finite element method designed for the solution of the system of time-domain Maxwell equations coupled to a Drude model for the dispersion of noble metals at optical frequencies is formulated on a tetrahedral mesh.

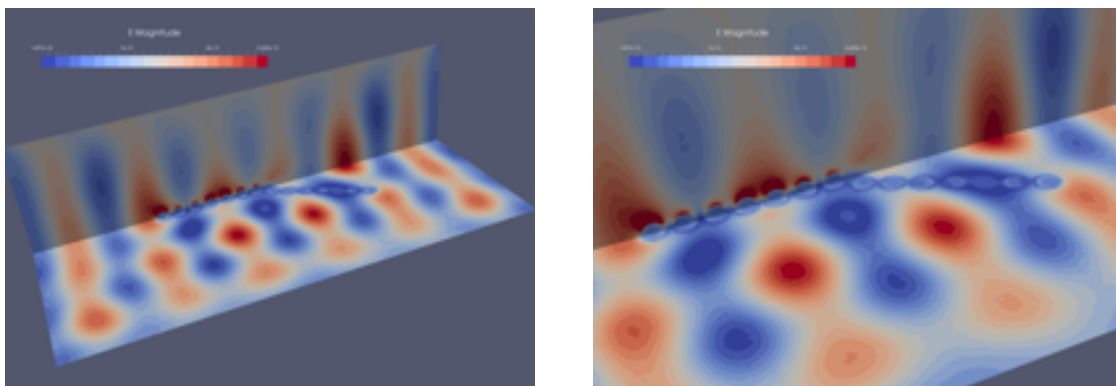


Figure 7. Y-shaped waveguide: contour lines of the amplitude of the discrete Fourier transform of the electric field.

6.4. Applications

6.4.1. Light diffusion in nanostructured optical fibers

Participants: Wilfried Blanc [Optical Fibers team, LPMC, Nice], Stéphane Lanteri, Paul Loriot, Claire Scheid.

Optical fibers are the basis for applications that have grown considerably in recent years (telecommunications, sensors, fiber lasers, etc.). Despite these undeniable successes, it is necessary to develop new generations of amplifying optical fibers that will overcome some limitations typical of silica. In this sense, the amplifying Transparent Glass Ceramics (TGC), and particularly the fibers based on this technology, open new perspectives that combine the mechanical and chemical properties of a glass host and the augmented spectroscopic properties of embedded nanoparticles, particularly rare earth-doped oxide nanoparticles. Such rare earth-doped silica-based optical fibers with transparent glass ceramic (TGC) core are fabricated by the Optical Fibers team of the Laboratory of Condensed Matter (LPMC) in Nice. The objective of this collaboration with Wilfried Blanc at LPMC is the study of optical transmission terms of loss due to scattering through the numerical simulation of light propagation in a nanostructured optical fiber core using a high order DGTD method developed in the team.

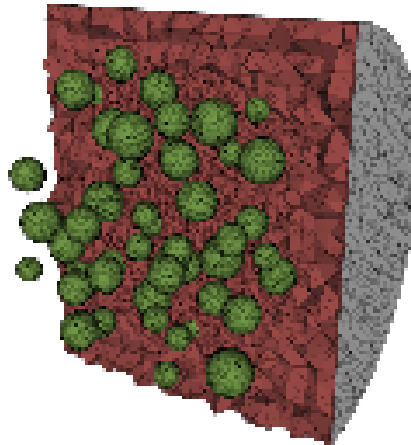


Figure 8. Unstructured tetrahedral mesh of a nanostructured optical fiber core.

6.4.2. Gap-plasmon confinement with gold nanocubes

Participants: Stéphane Lanteri, Antoine Moreau [Institut Pascal, Université Blaise Pascal], Claire Scheid, Jonathan Viquerat.

The propagation of light in a slit between metals is known to give rise to guided modes. When the slit is of nanometric size, plasmonic effects must be taken into account, since most of the mode propagates inside the metal. Indeed, light experiences an important slowing-down in the slit, the resulting mode being called *gap-plasmon*. Hence, a metallic structure presenting a nanometric slit can act as a light trap, i.e. light will accumulate in a reduced space and lead to very intense, localized fields. Recently, the chemical production of random arrangements of nanocubes on gold films at low cost was proved possible by Antoine Moreau and colleagues at Institut Pascal. Nanocubes are separated from the gold substrate by a dielectric spacer of variable thickness, thus forming a narrow slit under the cube. When excited from above, this configuration is able to support gap-plasmon modes which, once trapped, will keep bouncing back and forth inside the cavity. At visible frequencies, the lossy behavior of metals will cause the progressive absorption of the trapped electromagnetic field, turning the metallic nanocubes into efficient absorbers. The frequencies at which this

absorption occurs can be tuned by adjusting the dimensions of the nanocube and the spacer. In collaboration with Antoine Moreau, we propose to study numerically the impact of the geometric parameters of the problem on the behaviour of a single nanocube placed over a metallic slab (see Fig. 9). The behavior of single nanocubes on metallic plates has been simulated, for lateral sizes c ranging from 50 to 80 nm, and spacer thicknesses δ from 3 to 22 nm. The absorption efficiency in the cube Q_{cube} at the resonance frequency is retrieved from the results of each computation (see Fig. 10).

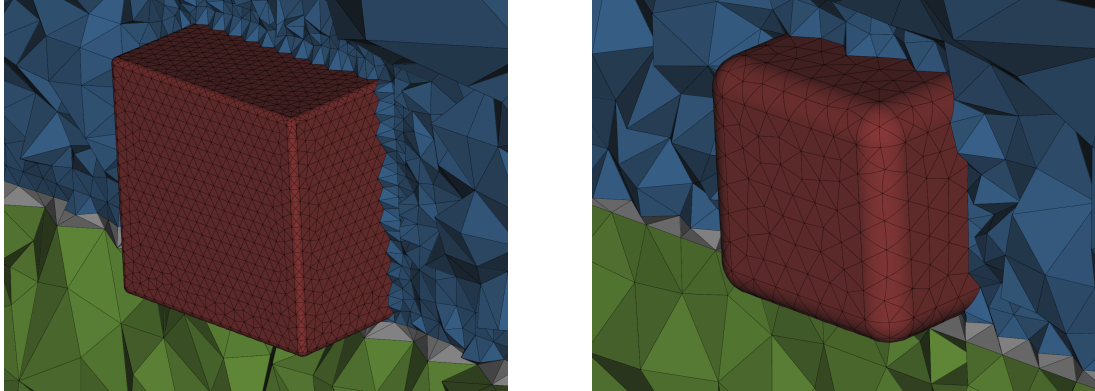
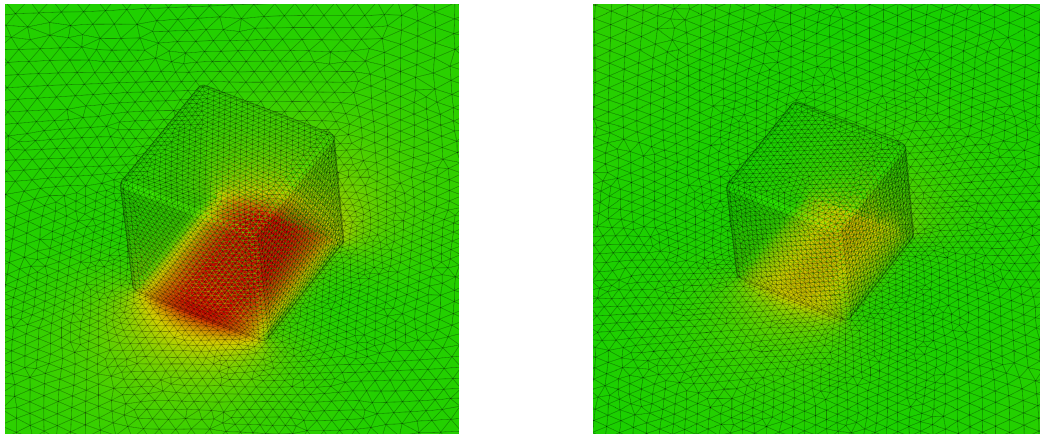


Figure 9. Meshes of rounded nanocubes with rounding radii ranging from 2 to 10 nm. Red cells correspond to the cube. The latter lies on the dielectric spacer (gray cells) and the metallic plate (green). Blue cells represent the air surrounding the device.



$$c = 70 \text{ nm}, \delta = 12 \text{ nm}$$

$$c = 60 \text{ nm}, \delta = 18 \text{ nm}$$

Figure 10. Amplitude of the discrete Fourier transform of the magnetic field for different nanocube configurations. All field maps are scaled identically for better comparison. The obtained field is more intense for configurations that yield high Q_{cube} values.

6.4.3. Light propagation in power splitters

Participants: Julien Coulet, Carlos Henrique Da Silva Santos [Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Brazil], Hugo Enrique Hernandez Figueroa [Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, São Paulo, Brazil], Stéphane Lanteri, Frédéric Valentin [LNCC, Petropolis, Brazil].

Power splitters are passive devices widely used in signal processing, which splits an input signal into two or more output signals. The repartition of the input power over each output is specific to the required usage. Even if power splitters are common in classical electronics, designing them at the micrometric scale is quietly recent and is an active field of research. The purpose of this study initiated in the framework of a collaboration with researchers at Unicamp in São Paulo is to study the electromagnetic wave propagation in such a power splitter geometry using a high order DGTD method developed in the team, see Fig. 11 .

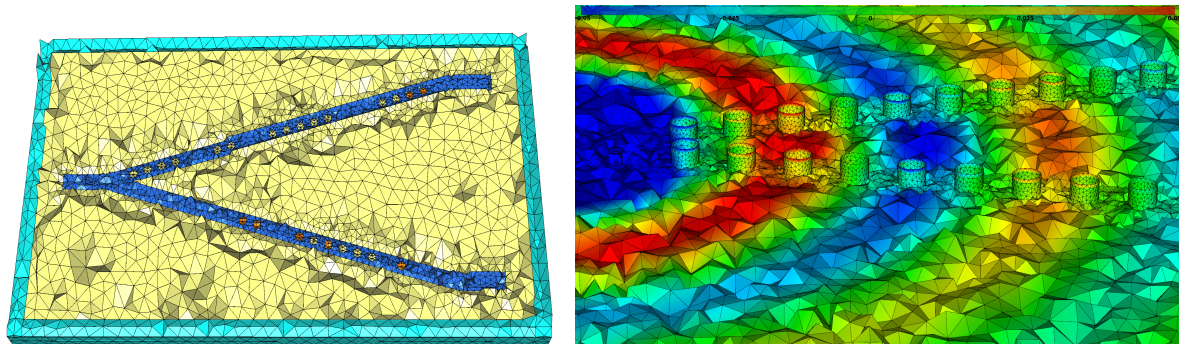
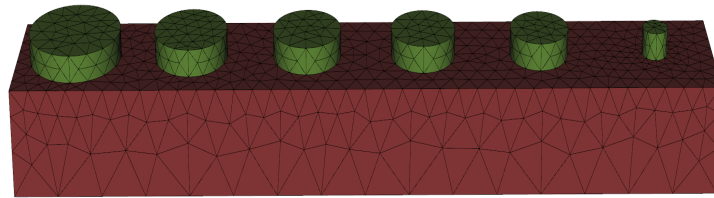


Figure 11. Geometry of Y-shaped power splitter (left) and contour lines of the amplitude of the electric field (right).

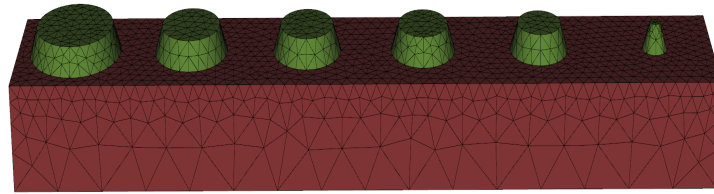
6.4.4. Dielectric reflectarrays

Participants: Maciej Klemm [Centre for Communications Research, University of Bristol], Stéphane Lanteri, Claire Scheid, Jonathan Viquerat.

In the past few years, important efforts have been deployed to find alternatives to on-chip, low-performance metal interconnects between devices. Because of the ever-increasing density of integrated components, intra- and inter-chip data communications have become a major bottleneck in the improvement of information processing. Given the compactness and the simple implantation of the devices, communications *via* free-space optics between nanoantenna-based arrays have recently drawn more attention. Here, we focus on a specific low-loss design of dielectric reflectarray (DRA), whose geometry is based on a periodic repartition of dielectric cylinders on a metallic plate. When illuminated in normal incidence, specific patterns of such resonators provide a constant phase gradient along the dielectric/metal interface, thus altering the phase of the incident wavefront. The gradient of phase shift generates an effective wavevector along the interface, which is able to deflect light from specular reflection. However, the flaws of the lithographic production process can lead to discrepancies between the ideal device and the actual resonator array. Here, we propose to exploit our DGTD solver to study the impact of the lithographic flaws on the performance of a 1D reflectarray (see Fig. 12). Efficient computations are obtained by combining high-order polynomial approximation with curvilinear meshing of the resonators, yielding accurate results on very coarse meshes (see Fig. 13). The study is continued with the computation of the reflection of a 2D reflectarray. This work constitutes the base of a wider study in collaboration with Maciej Klemm at the Centre for Communications Research, University of Bristol.

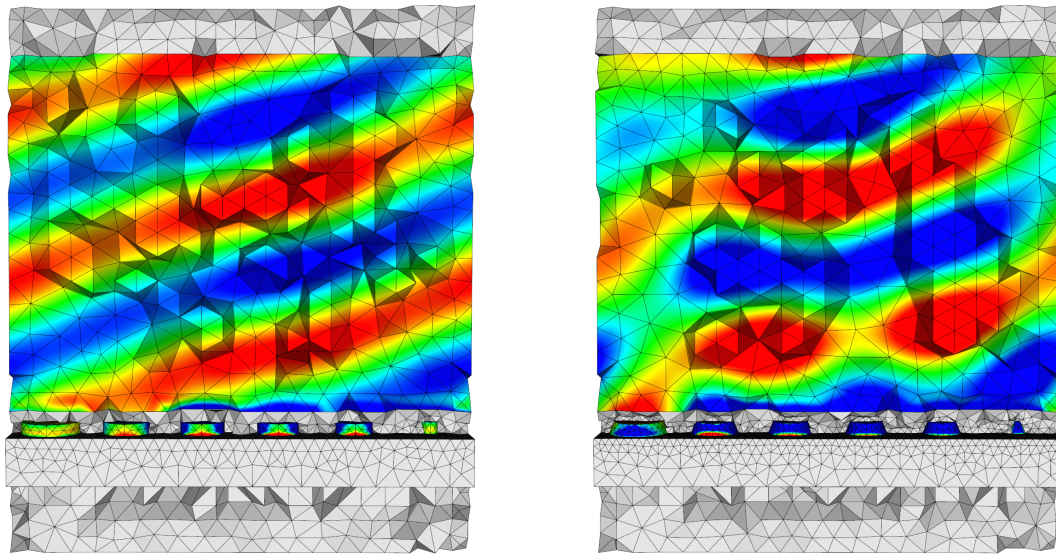


Ideal reflectarray



Realistic reflectarray

Figure 12. Ideal and realistic 1D dielectric reflectarray meshes. The red tetrahedra correspond to silver, while the green ones are made of an anisotropic dielectric material. The device is surrounded by air and terminated by a PML above and below, and by periodic boundary conditions on the lateral sides.



Ideal reflectarray

Realistic reflectarray

Figure 13. Time-domain snapshot of E_y component for ideal and realistic 1D dielectric reflectarrays. Solution is obtained in established regime at $t = 0.1$ ps. Fields are scaled to $[-1, 1]$.

NANO-D Project-Team

7. New Results

7.1. Algorithms for Orbital-Free Density Functional Theory

Participants: Francois Rouse, Stephane Redon.

The Schrödinger equation permits, in theory, to model and simulate every molecular systems exactly. Unfortunately it is not computationally doable to solve this equation even on really small systems (2 atoms). Density Functional Theory (DFT) gives a method to solve this equation, find the electronic structure and simulate molecules with the laws of physics on reasonably large system : from 1.000 to 10.000 depending on the basis chosen and the version of DFT used. Unfortunately, the computation of kinetic energy requires the orthogonalization of the basis, which consumes a lot of time and prevents the algorithm from being adaptive : one needs to recompute the whole system if a little change is done in the molecules position. One can deals with this issue by computing the kinetic energy directly with the electronic density and not anymore with the orbitals. That is the idea of Orbital-Free DFT (OF-DFT). It can models great systems (up to 1.000.000 atoms) and be turned adaptive. On the other hand, it looses a lot of accuracy and power to model different kind of systems on the other DFT.

We have already developed our own OF-DFT code. It runs on parallel cores, is implemented in the SAMSON platform as a SAMSON App and gives correct electron's densities. The electronic structures are computed in real space to preserve the possibility of incremental calculations. We are now going to test our implementation, and will then attempt to make the method adaptive. The difficulty will be the determination of the domain that needs to be recomputed when a part of the system has moved, and the criteria that will help to do so.

7.2. Parallel adaptively restrained particle simulations

Participants: Krishna Kant Singh, Stephane Redon.

We have continued our work on the development of parallel adaptively restrained particle simulations. We developed new algorithms for neighbor list and incremental force updates. These algorithms have advantages over the state-of-the-art methods for simulating a system using Adaptively Restrained Molecular Dynamics (ARMD). We have simulated systems with different number (500, 4000 and 108000) of LJ particles using adaptively restrained integrator and Lennard-Jones potential in NVE (constant number of particles, Volume and Energy) and NVT ensemble (constant number of particles, Volume and Temperature). All the particles were placed in an orthogonal box. We used periodic periodic boundary conditions with 8.5 angstrom cut-off for the Lennard-Jones potential. The system was simulated using 2 femtoseconds. We compared the LAMMPS algorithm to adaptive algorithm while using adaptively restrained integrator. Our results show that a significant speed-up can be achieved if more than 60% of the particles are restrained (Figure 4). Figure 5 shows that ARMD in NVT ensemble preserves the average temperature of the system (irrespective of number of restrained particles).

7.3. Incremental algorithms for long-range interactions

Participants: Semehor Etorh, Stephane Redon.

Numerical simulations of molecular dynamics (MD) are very expensive in terms of CPU resources. During Molecular dynamics simulations, the most CPU intensive task is the evaluation of the interaction potential [78]. Due to the large number of particles involved, updating this potential may have, at each time-step, a very high computational cost.

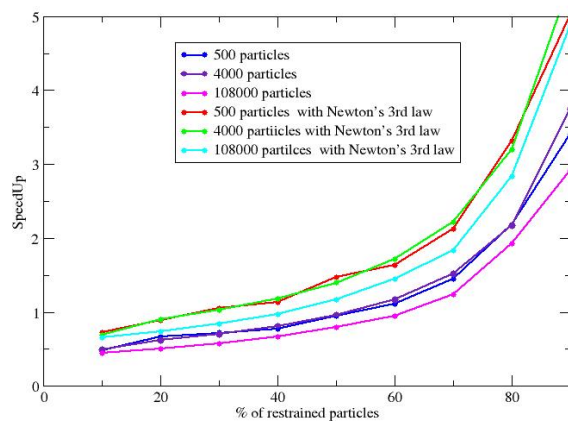


Figure 4. Speedup using ARMD on different benchmark

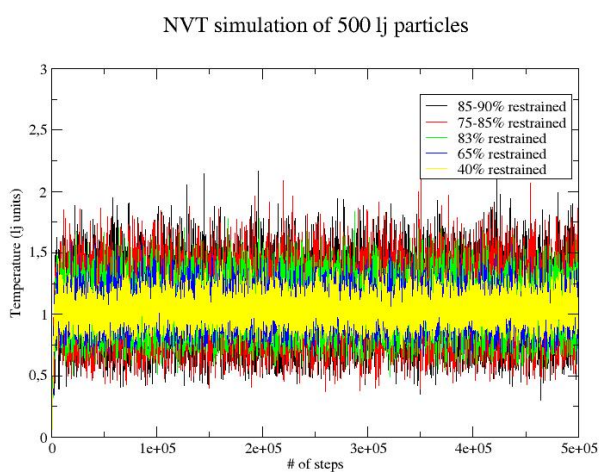


Figure 5. Temperature profile of 500 LJ particles in NVT ensemble using ARMD

In large crystalline ionic system, Ewald summation is the most popular method for computing electrostatic interactions. It rewrites the interaction potential ϕ as the sum of a short-range term and a long-range term. Ewald summation using optimal parameters requires $\mathcal{O}(N^{3/2})$ operations [47], [30] but it can be modified so that it involves only about $\mathcal{O}(N \log N)$ operations [31], [85] by using the Fast Fourier Transform.

We want to develop a new approach that can reduce the computational cost by using incremental algorithms. The key-idea is to use, for each time-step of the simulation, information that we have computed in previous steps.

The Particle Mesh Ewald (PME) algorithm developed by Darden et al. is the most successful approach for computing long range interactions. In the particle mesh method, just as in standard Ewald summation, the generic interaction potential is separated into two terms. The so-called short-range contribution can be easily calculated in a direct space by using truncation methods. Where as the long-range contribution is calculated using two Fast Fourier transforms ($N \log(N)$ algorithm). In practice, the long range contribution algorithm boil than to [30]:

- Map particle charge density Q to a mesh
- Compute the forward Fast Fourier Transform of the approximation Q_m of charge density on the mesh
- Multiply Q_m by a green function (related to the choice of the mesh).
- Compute a backward Fast Fourier Transform of the result.
- Retrieve the long-range contribution potential by interpolating the previous result at particles positions.

We modified this algorithm to make it incremental. We started from the PME implementation in LAMMPS. Instead of mapping the charge density to the mesh, we mapped the increment of density dQ to the mesh. The FFT solver KissFFT is based on a divide-and-conquer algorithm. We built a sparse input solver as a modified version of FFT solver which computes only needed (non trivial) operations [74]. We built also a sparse output solver inspired by the algorithm proposed by Katabi et al. [36]. Unfortunately, we did not get significant speed-ups with these modifications.

We decided to compute the increment of the long-range contribution related to the increment of density dQ by using multi-resolution methods. These methods are slower than PME but have better adaptive behavior. The multigrid approach was chosen because of its $\mathcal{O}(N)$ behavior and its good scalability [13]. We are currently developing an adaptive multigrid method.

7.4. Motion planning architecture for nanosystems

Participants: Leonard Jaillet, Stephane Redon.

In the past, we have started the development of original quasi-static simulation methods for nano-scale systems, based on motion planning methods inspired from Robotics. In the continuity of this work, we have proposed an original Motion Planning architecture for nanosystems platform called planning. This platform offers a general framework for motion planning applied to nanosystems. In particular it includes:

- A flexible definition of the degrees of freedom that describe the system, allowing different levels of representation (e.g. Cartesian coordinates, internal coordinates, coarse grain representation, etc.).
- The possibility to define an arbitrary set of initial, final and intermediate states, guiding the search for a solution path.
- The possibility to define an arbitrary set of constraints on the intermediate states of the path (e.g. geometric constraints, energy constraints, etc.)
- Several modular functionalities specific to motion planning (e.g. conformational sampling, exploration strategy, nearest neighbor search, etc.)
- An adapted integration within SAMSON which allows using directly all the existing force fields and state updaters present in the platform.

The planning architecture has been the base of several SAMSON modules. In particular, it led to the Planner-Explore module, which regroups many of the functionalities proposed and that can be combined together through a graphical interface. This module has in particular been used to study two complex problems:

- To capture the transition paths between endiandric acids (see Figure 6).
- To find the global minima of Lennard-Jones clusters, for dimensions up to one hundred.

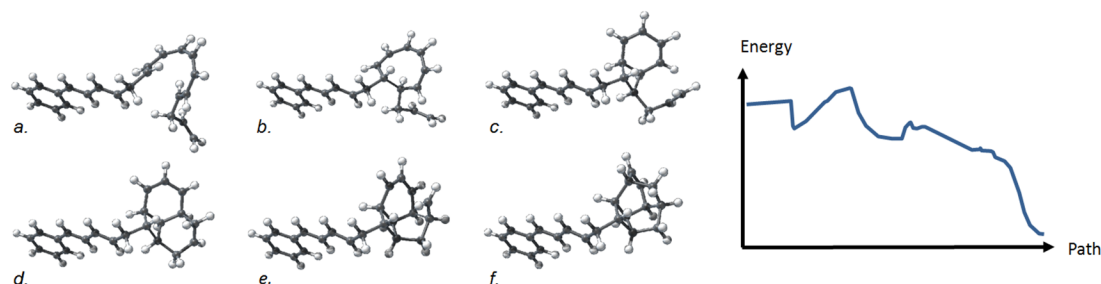


Figure 6. Transition path with its corresponding energy for an homolog of the endiandric acid and produced thanks to the Planner-Explore module.

7.5. Optimization of transition paths

Participants: Leonard Jaillet, Stephane Redon.

Motion planning methods allow producing initial paths which represent transitions from one given conformation to another. However, these paths are typically suboptimal because of the probabilistic nature of the search strategy. Hence, it is necessary to develop tools to locally increase the path quality of the solution generated during the first phase. We have proposed several methods to address such a problem. One method developed is a variant of a state-of-the-art approach called nudged elastic band (NEB). It optimizes a set of intermediate images along the path, such that each image finds the lowest energy possible while maintaining equal spacing to neighboring images. Another technique we proposed is to rely on an equivalent of the shortcutting technique developed in Robotics motion planning, but applied to the context of energy landscape. Finally, we also have complemented these methods with additional tools to do simple path edition such as cutting or thinning paths.

7.6. As-rigid-as-possible shape interpolation for molecular modeling

Participants: Minh Khoa Nguyen, Leonard Jaillet, Stephane Redon.

Computer-aided methods play an important role in the study of molecular structures and interactions. Inspired by the as-rigid-as-possible approaches in the field of computer graphics, we created a tool for studying large deformation of molecular structures. This tool generates interpolated structures between two known conformations of a molecule while satisfying physical constraints. The users may use it for exploring, preprocessing, or combining their model with other biological algorithms. The developed method is flexible and can be extended to include physical properties of molecular structures.

We tested our method on a graphene sheet folding into a nanotube (Figure 8) and a few biological molecules, one of which is shown in Figure 9. The results show realistic transition motions compared to those from the linear interpolation approach.

The ARAP interpolation method has two main advantages: simplicity and preservation of local rigidity. The method is totally geometrical, yet can be extended to include physical or biological properties such as bond strength. It will be proposed as a SAMSON Element for the SAMSON software platform for computational nanoscience.

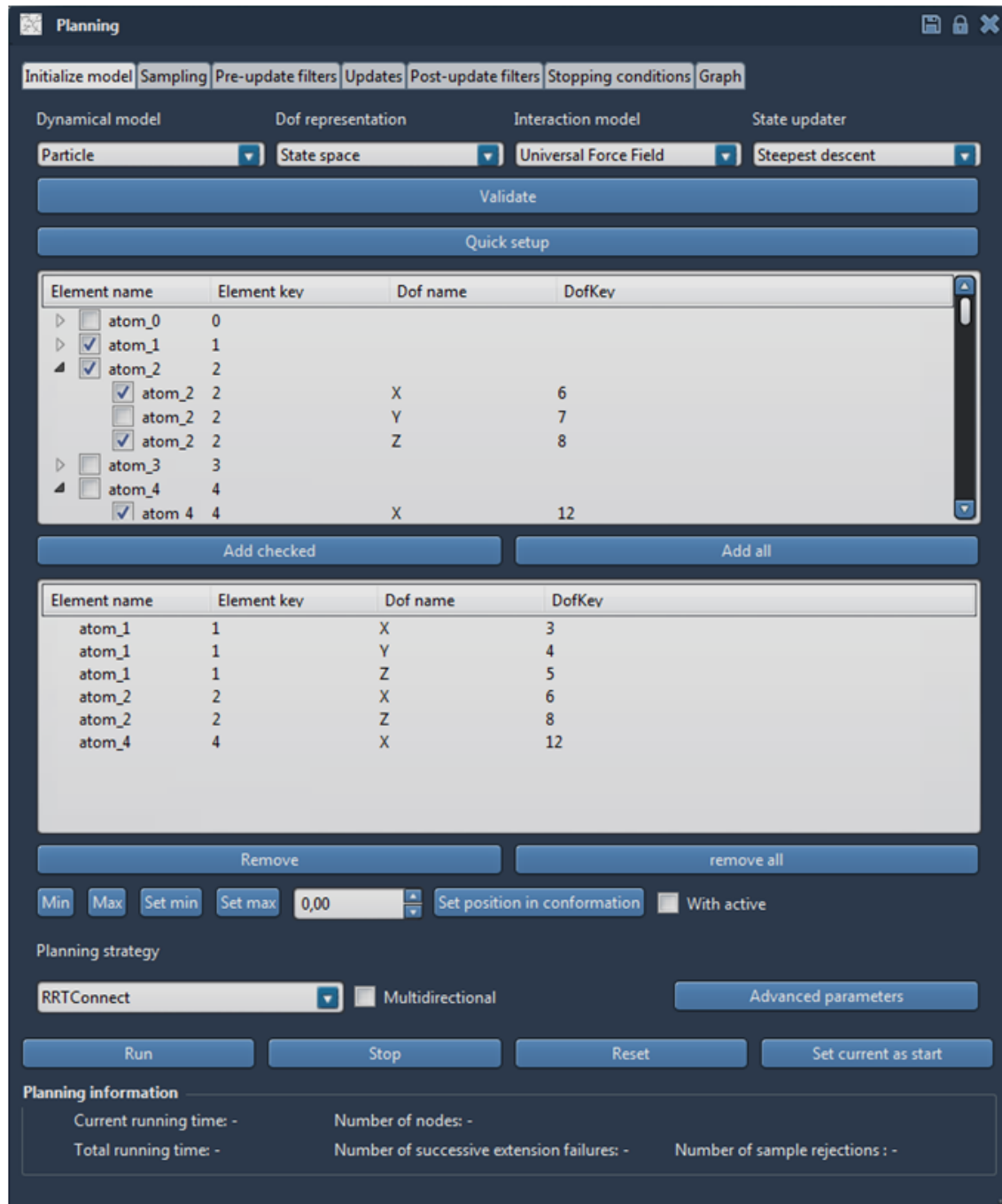


Figure 7. One tab of the graphical interface of the Planner-Explore modules, which allows initializing the model to be simulated.

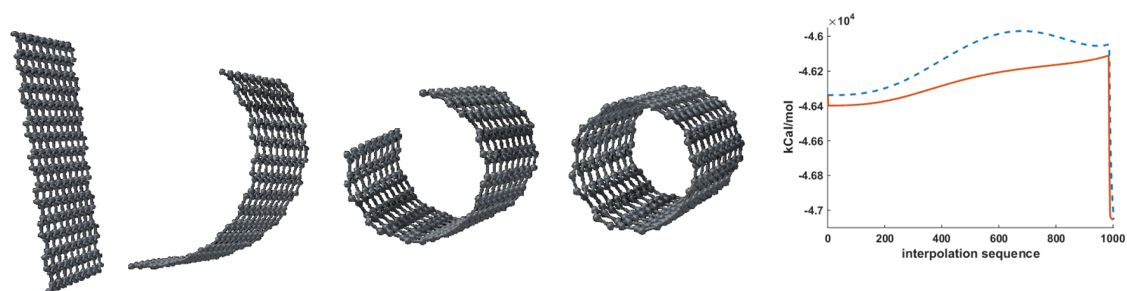


Figure 8. ARAP interpolation to generate graphene sheet folding into a nanotube. The last figure plots the energy for a sequence of 1000 interpolated images. The energy of ARAP interpolation is shown by the dotted blue curve and the optimal energy after applying NEB is shown by the solid red curve.

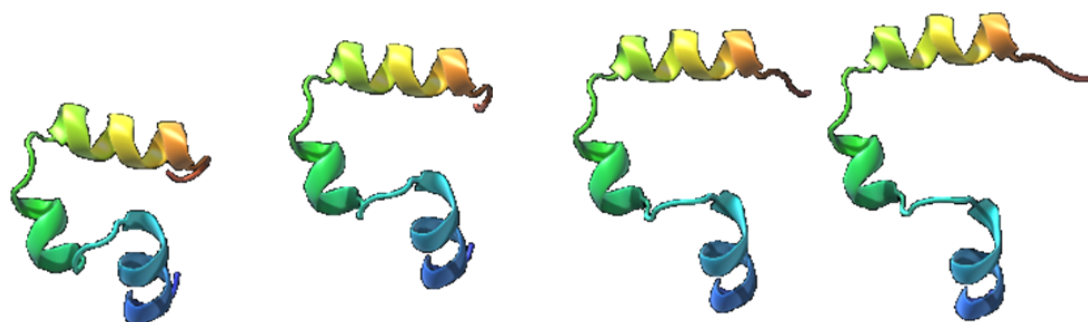


Figure 9. Transition obtained by the ARAP method of a subdomain of the villin headpiece (protein ID: 1YFR) into its distorted shape generated manually.

7.7. Automatic parameterization for the Universal Force Field

Participants: Svetlana Artemova, Leonard Jaillet, Stephane Redon.

We have continued working on the integration of the Universal Force Field in SAMSON. This force field is a classical non-reactive force field that has parameterizations for all atoms of the periodic table with atomic number lower than 103. Our implementation of this force field includes a new automatic perception scheme for molecular systems that is specifically-tailored for UFF, as well as several corrections and refinements that have been lately proposed in the literature. We have tested this implementation on more benchmarks and improved its computational performance. Additionally, we have compared our implementation to that of the OpenBabel toolbox. As a result, our self-contained implementation was integrated in a new module for SAMSON and is now available on SAMSON-Connect website (see Figure 10). The paper describing the obtained results will appear in the Journal of Computational Chemistry.

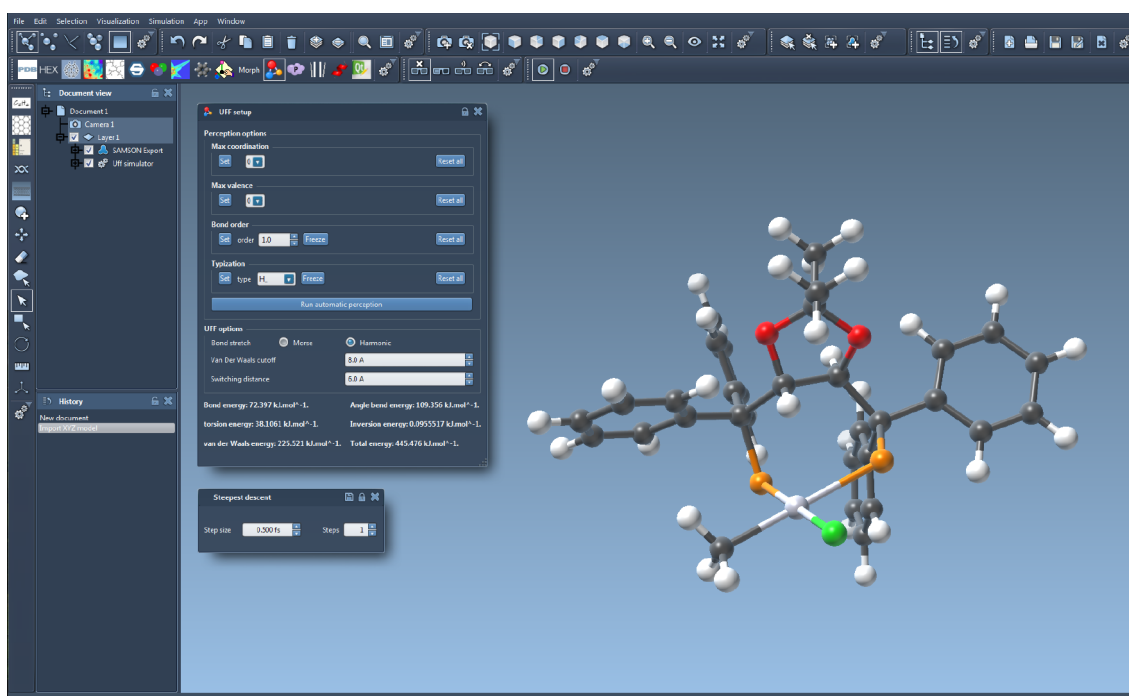


Figure 10. A molecule being interactively manipulated in SAMSON thanks to the UFF module. The interface of the UFF module allows to setup UFF. The upper part of the interface proposes options to manually adjust the perception of the molecular system. The middle part proposes the UFF options. The lower part prints out each energy contribution with the resulting total energy.

7.8. Interactive modeling with the Universal Force Field

Participants: Leonard Jaillet, Svetlana Artemova, Stephane Redon.

In parallel with the classical Universal Force Field, we have continued working on an extension of this force field that we call Interactive Modeling UFF (IM-UFF). In classical UFF topologies and atoms' typizations are set in the initialization phase and remain fixed for the entire simulation. IM-UFF, on the contrary, allows soft transitions for both topologies and atoms' typizations. This new approach, thus, combines the possibility to

significantly modify molecular structures (as with reactive force fields) with a broad diversity of supported systems thanks to the universality of UFF. Such an extension lets the user easily build and edit molecular systems interactively while being guided by physically-based inter-atomic forces. The validity of this extended version of UFF was tested on the same large set of benchmarks as those used to test classical UFF, and the results of both approaches were compared.

7.9. Error Analysis of Modified Langevin Dynamics

Participants: Zofia Trstanova, Gabriel Stoltz, Stephane Redon.

Adaptively Restrained Particles Simulations (ARPS) were recently proposed with the purpose of speeding up molecular simulations. The main idea is to modify the Hamiltonian such that the kinetic energy is set to zero for small velocities, which allows to save computational time since particles do not move and forces need not be updated. ARPS can be combined with Langevin dynamics in order to speed up the computation of macroscopic quantities.

The aim of this work is to understand how simulation errors depend on the parameters of the method. We distinguish the statistical error and the systematic error related to the finiteness of the time step Δt . The statistical error is controlled by variance, that is given by

$$\sigma^2 = -2 \langle A - \mu(A), \mathcal{L}^{-1}(A - \mu(A)) \rangle_{L^2(\mu)} \quad (7)$$

where μ is the invariant measure, \mathcal{L} is the generator of the stochastic process and A an observable. First we demonstrate by use of weighted L^∞ estimations that the ARPS-Langevin dynamics are well defined. In the main part of this work, we quantify the increase of variance of the ARPS-Langevin process as a function of the ARPS parameters. For small parameters, we express the generator of the ARPS-Langevin dynamics as a perturbed generator of the Langevin dynamics, and study the asymptotic expansions of the variance (1) in the restrained dynamics parameter ε .

$$\sigma_\varepsilon^2 = \sigma^2 + \mathcal{O}(\varepsilon)$$

For large values of ε , we perform numerical simulations. For a simple 1D system we approximate \mathcal{L}^{-1} by Galerkin approach and for higher-dimensional systems we discretize the stochastic differential equations by a second order method and analyze a model of a dimer surrounded by solvent particles.

7.10. Algorithmic speed up of the ARPS method

Participants: Zofia Trstanova, Gabriel Stoltz, Stephane Redon.

Adaptively Restrained Particles Simulations (ARPS) allow to save computational time at each time step since particles do not move and forces need not be updated. The associated gain can be quantified by an algorithmic speed-up factor $S_{\text{algo}} \geq 1$. Intuitively, freezing more particles leads to larger algorithmic speed-ups, but also larger correlations in time.

We analyzed the algorithmic speed up with respect to the standard methods. Since the ARPS algorithm is based on adding and subtracting of the forces between active particles, the gain with respect to the standard method, where only one complete computation of all interactions is performed at each time step, is achieved only if the percentage of restrained particles is big enough. Hence we studied the necessary conditions, under which the computational complexity of the forces updating in the ARPS method is lower than the one of the standard method. This allows to achieve an algorithmic speed up that is always bigger than one.

We also propose a simple strategy for choosing optimal simulation parameters.

7.11. Numerical analysis for the ARPS method

Participants: Zofia Trstanova, Gabriel Stoltz, Stephane Redon.

Previous works have led to understanding of the choice of optimal parameters for the ARPS dynamics. The interest lies in achieving the highest percentage of restrained particles, while minimizing the modification of the variance and the systematic error. We study discretization schemes of the ARPS-Langevin dynamics, such that the systematic error remains of second order in the time step size and we introduce a Metropolis step in order to stabilize the simulations and hence to allow "a sharper" choice of the ARPS parameters, which lead to better algorithmic speed-ups.

7.12. New rendering algorithm for secondary structures

Participants: Marc Aubert, Stephane Redon.

We developed a new algorithm for rendering secondary structures of proteins (Figure 11). The method relies on the determination of the most probable secondary structure elements (e.g. alpha helices and beta sheets) based on geometrical features of a protein. After construction of control points on the CPU, the method generates triangles directly on the Graphics Processing Unit (GPU) through geometry shaders. The number of generated triangles may be adaptively chosen based on e.g. the camera distance and the desired resolution. The secondary structure algorithm and the rendering algorithm are both fast enough to allow for interactive modification of the protein (e.g. thanks to As-Rigid-As-Possible editing algorithms).

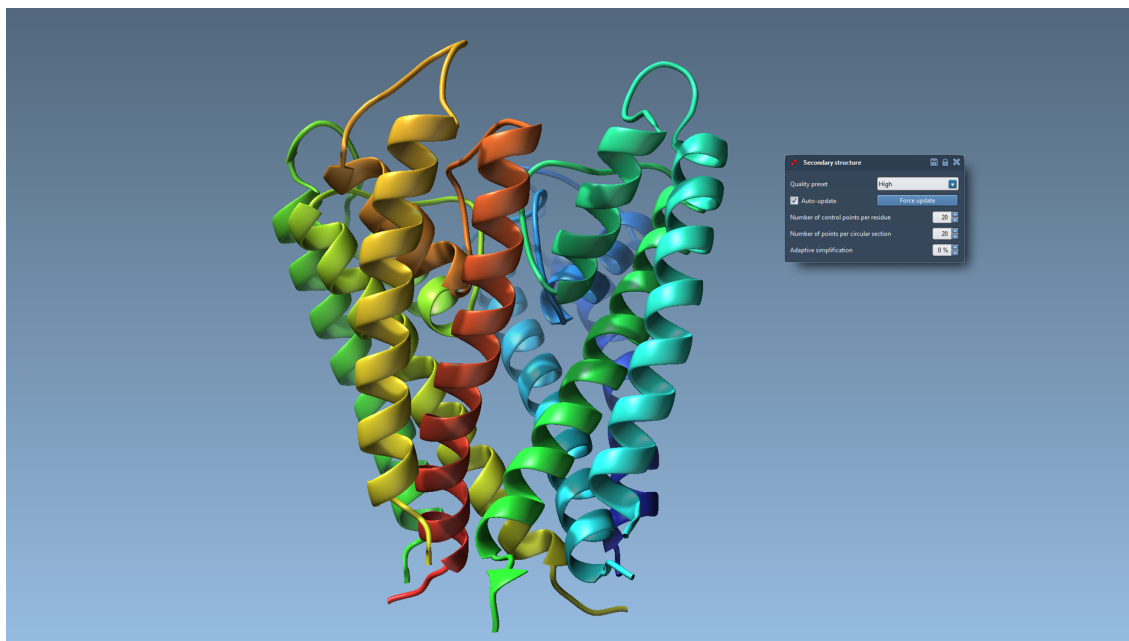


Figure 11. Protein secondary structure rendering on GPUs.

7.13. Property models

Participants: Marc Aubert, Stephane Redon.

We extended the hierarchy of classes in SAMSON for *property models*. Property models are one of the five categories of models in SAMSON, with structural models (for geometry and topology), visual models (for custom graphical representations), dynamical models (to describe degrees of freedom) and interaction models (to represent energies and forces). We have added classes to easily represent in SAMSON various functions, fields (e.g. scalar fields and vector fields), etc. These property models are template classes which may rely on the unit system of SAMSON to perform dimensional analysis at compile time.

7.14. Integration of tools in SAMSON

Participants: Nadhir Abdellatif, Stephane Redon.

Thanks to funding from the Nanosciences Foundation in Grenoble, we developed SAMSON Elements (modules for SAMSON) that integrate existing tools. In particular, we integrated OpenBabel, a tool to convert between numerous molecular formats (Figure 12), ClustalW, a tool for sequence alignment (Figure 13), and Pepsi-SAXS, a tool for SAXS developed in the group (Figure 14).

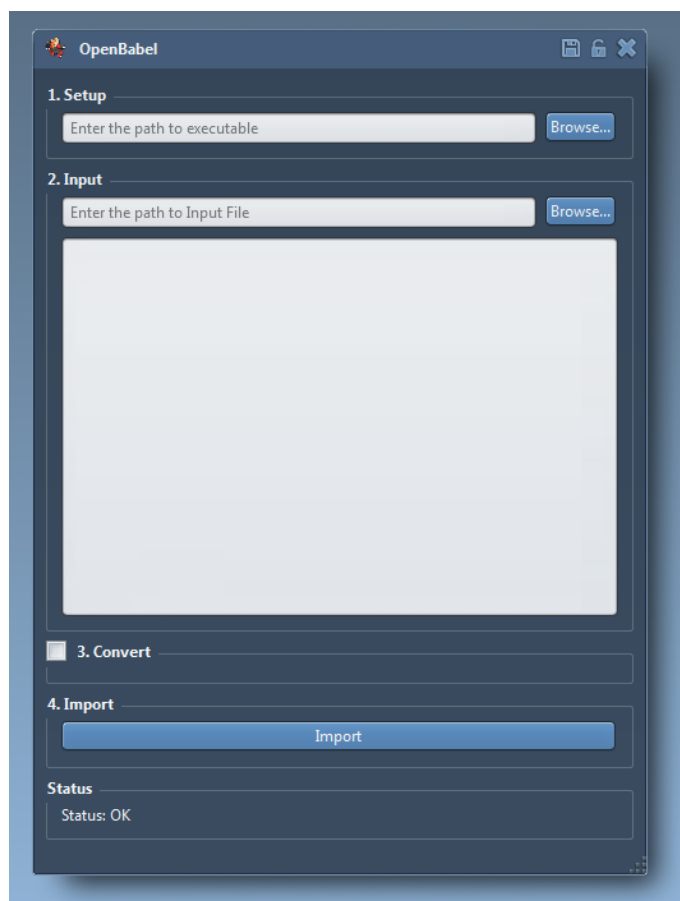


Figure 12. The OpenBabel connector in SAMSON

7.15. Development of SAMSON Connect

Participants: Mohamed Yengui, Jocelyn Gate, Stephane Redon.

We have continued the development of SAMSON Connect (Figure 15, <https://www.samson-connect.net>), the online platform for distributing SAMSON and SAMSON Elements (modules for SAMSON). SAMSON Connect is a web application, associated to a database, that functions as the well-known stores for mobile Apps (e.g. Google Play, the Apple App store, etc.). Users may create an account, download SAMSON, and add SAMSON Elements to their configuration based on their needs (Figure 16). Adding a SAMSON Element is performed in just one click (to the Add button of the corresponding SAMSON Element), and the SAMSON



Figure 13. ClustalW in SAMSON

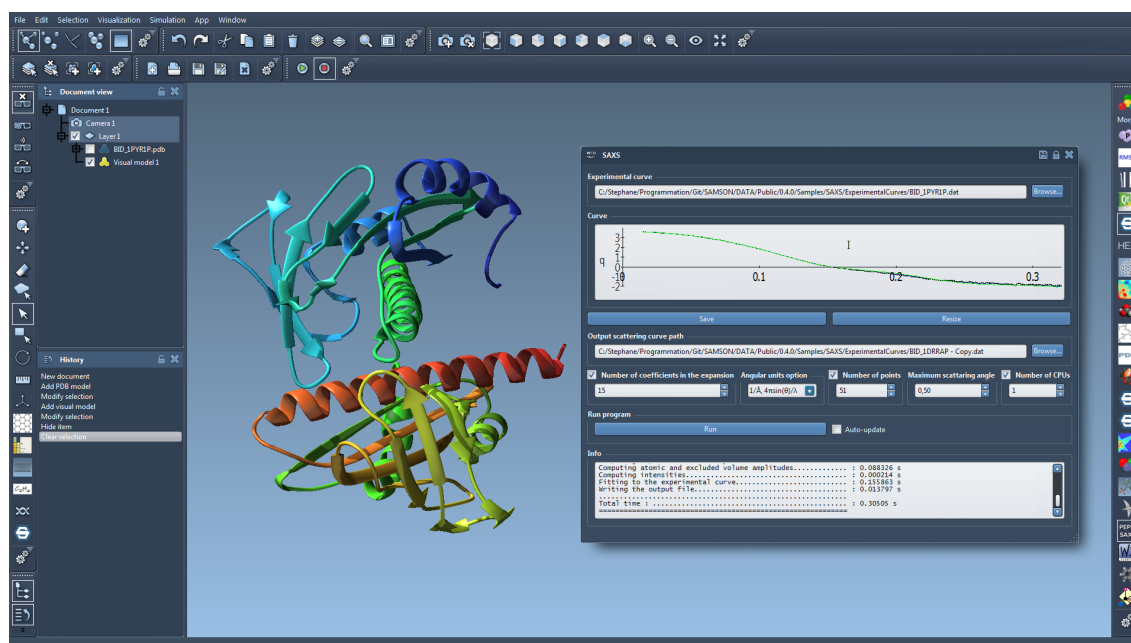


Figure 14. Our Pepsi-SAXS tool integrated in SAMSON

Element is installed when the user restarts SAMSON. Users may also request an upgrade to a Developer status, after which they can download the SAMSON Software Development Kit used to develop SAMSON Elements. They may then upload their SAMSON Elements to SAMSON Connect in order to share them. The platform opened in March 2015 to release the first beta version of SAMSON. We also produced some video tutorials for SAMSON (Figure 17).

On the back-office of SAMSON Connect, we added several functionalities that facilitate publishing new versions of SAMSON and SAMSON Elements (e.g. choosing default SAMSON Elements), email users based on their account type (user, developer, etc.), OS, etc. We also turned to automatic acceptance of new user accounts (once they validate their email address). We also updated the SAMSON web service to enable more message types and retrieve information about the server, the database, etc.

7.16. Documenting the SAMSON Software Development Kit

Participants: Stephane Redon, Jocelyn Gate, Svetlana Artemova.

The SAMSON Software Development Kit (SDK) is at the core of the SAMSON platform and makes it possible to develop SAMSON Elements (modules). The API of SAMSON contains numerous classes and allows for a variety of modules types (e.g. parsers, force fields, visual models, integrators, apps, editors, etc.), and provides several non-elementary mechanisms (e.g. a unit system, a signals and slots mechanism, memory management, data structures for incremental calculations, etc.). We continued writing the SDK documentation accessible to SAMSON Elements developers. The current PDF version for the beta 0.4.0 version has passed 500 pages.

7.17. SAMSON SDK Helpers

Participants: Jocelyn Gate, Stephane Redon.

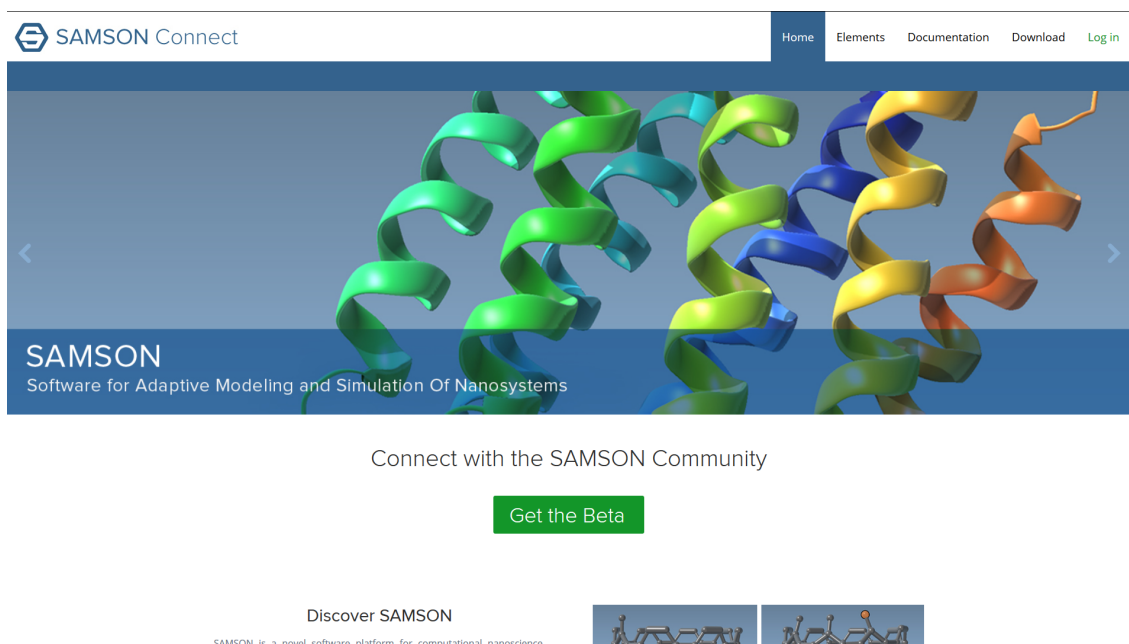


Figure 15. The home page of SAMSON Connect (<https://www.samson-connect.net>)

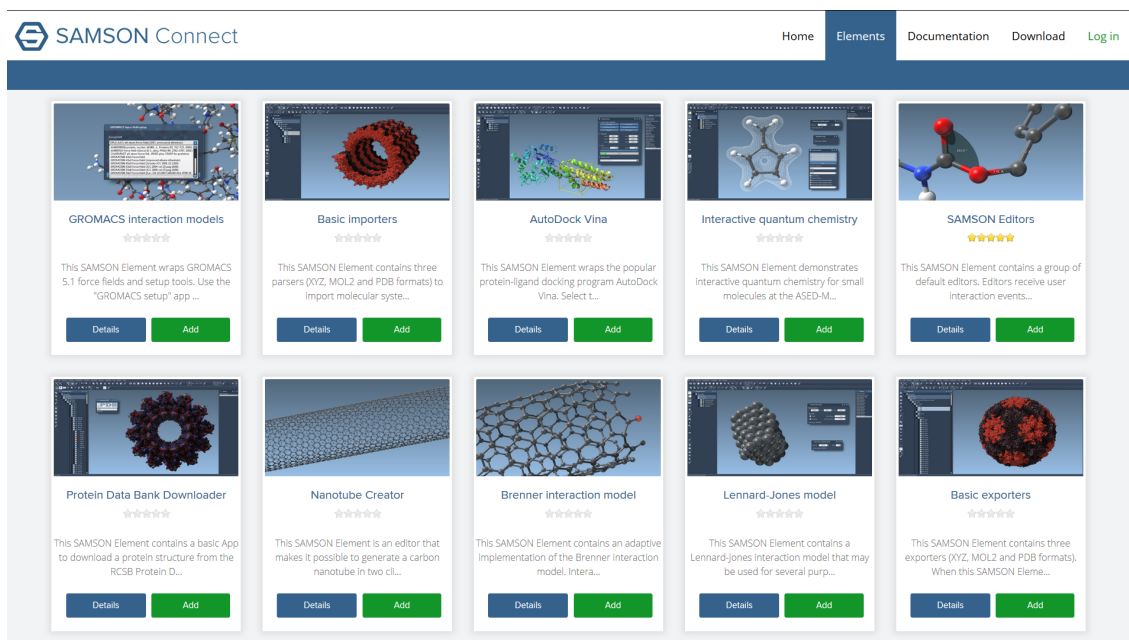


Figure 16. The Elements page, where users may add SAMSON Elements (modules) to their configuration (<https://www.samson-connect.net>)

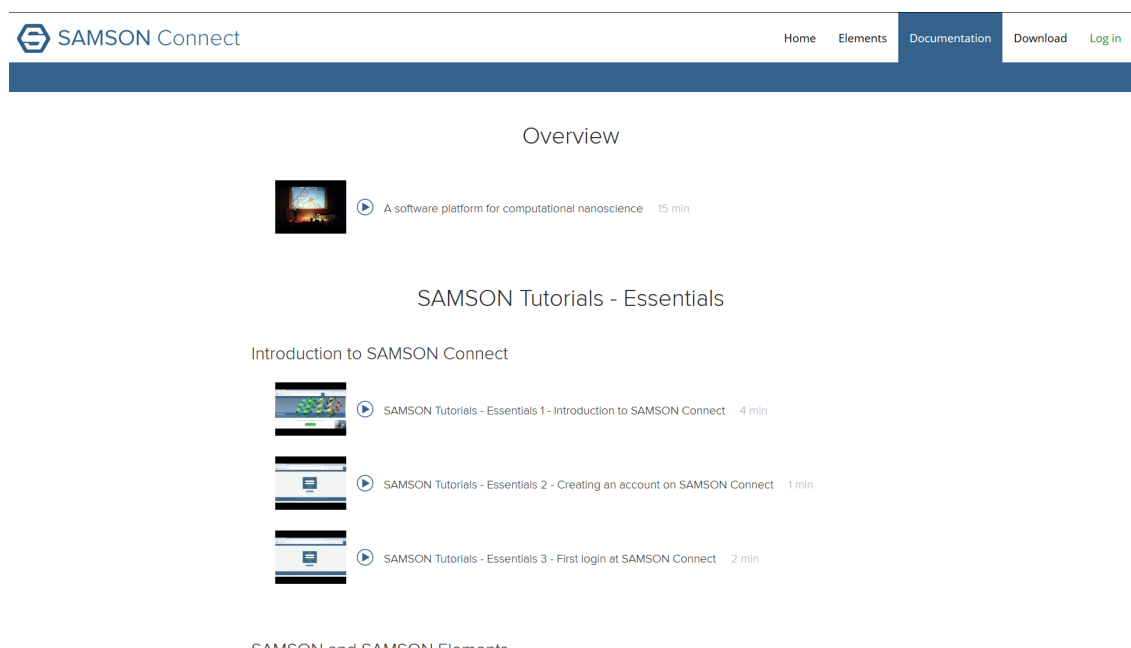


Figure 17. The documentation page on SAMSON Connect (<https://www.samson-connect.net>)

We have developed a number of Helpers in the SAMSON SDK, in order to facilitate the development of SAMSON Elements (modules for SAMSON). For example, the SAMSON Element generator (Figure 18) generates code that immediately compiles and runs, and that developers may complete, for a number of SAMSON Classes.

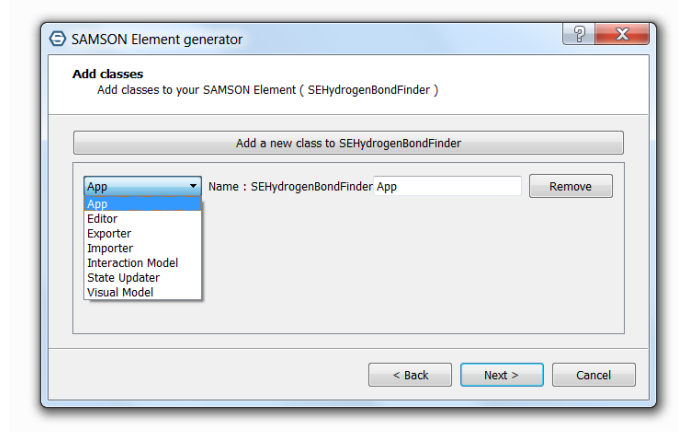


Figure 18. The SAMSON Element Generator makes it easy to develop modules for SAMSON

We have also developed for the group a helper able to upload numerous SAMSON Elements to SAMSON Connect at the same time (Figure 19), which is especially useful given the rapidly growing number of modules being developed in the team.

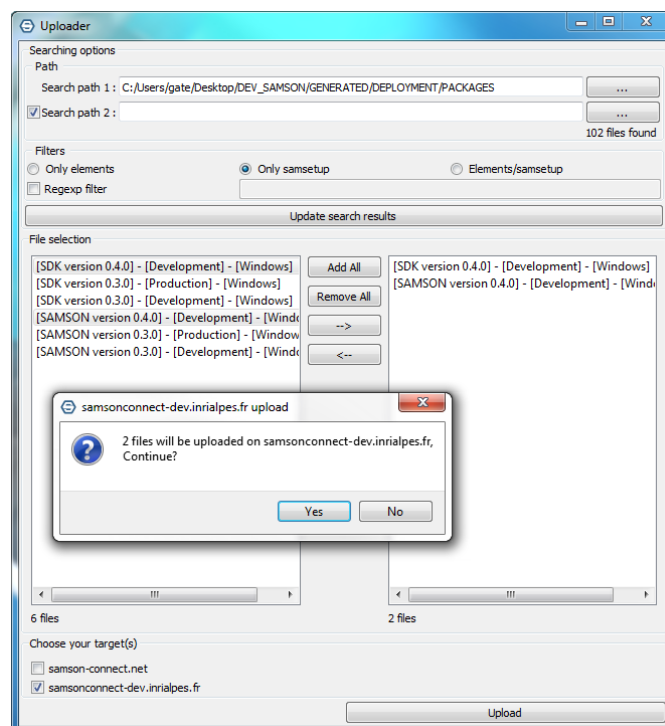


Figure 19. The SAMSON Element uploader eases the transfer of multiple SAMSON Elements to SAMSON Connect

7.18. Pepsi-SAXS : an adaptive method for rapid and accurate computation of small angle X-ray scattering profiles

Participant: Sergei Grudinin.

We developed a new method called Pepsi-SAXS that calculates small angle X-ray scattering profiles from atomistic models. Our method is based on the multipole expansion scheme and is significantly faster and more precise compared to other tested methods. In particular, using the Nyquist-Shannon-Kotelnikov sampling theorem, we adapt the multipole expansion order to the size of the model and the resolution of the experimental data. We argue that using the adaptive expansion order, our method has the same quadratic dependence on the number of atoms in the model as the Debye-based approach, however, with a much smaller prefactor in the computational complexity.

We have systematically validated our method on an excessive set of over fifty models collected from the BioIsis and SASBDB databases. Using a laptop, we demonstrated that Pepsi-SAXS is about 9, 33 and 43 times faster compared to CRY SOL, FoXS and the 3D-Zernike method in SAS t b x, correspondingly, when tested on data from the BioIsis database, and is about 5, 18 and 23 times faster compared to CRY SOL, FoXS and SAS t b x, correspondingly, when tested on data from SASBDB. On average, Pepsi-SAXS achieves 17% smaller value of χ compared to CRY SOL and 15% smaller value of χ compared to FoXS for BioIsis profiles, and 6% smaller value of χ compared to CRY SOL and 19% smaller value of χ compared to FoXS for SASBDB profiles.

7.19. Knodle: a Support Vector Machines-based automatic perception of organic molecules from 3D coordinates

Participants: Maria Kadukova, Sergei Grudinin.

We addressed the problem of the assignment of atom types and bond orders in low molecular weight compounds. For this purpose, we developed a prediction model based on nonlinear Support Vector Machines (SVM), implemented in a KNOWledge-Driven Ligand Extractor called *Knodle*, a software library for the recognition of atomic types, hybridization states and bond orders in the structures of small molecules. We trained the model using an excessive amount of structural data collected from the PDBbindCN database. Accuracy of the results and the running time of our method is comparable with other popular methods, such as NAOMI, fconf, and I-interpret. More precisely, on the popular Labute's benchmark set consisting of 179 protein-ligand complexes, *Knodle* makes five to six perception errors, NAOMI makes seven errors, I-interpret makes nine errors, and fconv makes thirteen errors. On a larger set of 3,000 protein-ligand structures collected from the PDBBindCN general data set (v2014), *Knodle* along with NAOMI have a comparable accuracy of approximately 6 % of errors, whereas fconv produces approximately 13 % of errors. Overall, our study demonstrates the efficiency of nonlinear SVM in structure perception tasks.

7.20. Symmetry Detection Method

Participants: Silvia Dias Pinto, Sergei Grudinin.

We developed an algorithm for automatic recognition of the point group symmetry in electron density maps of biological objects. More precisely, the method operates on cryo-Electron Microscopy (cryoEM) data, which typically contain 3D structures of multi-domain proteins and their complexes. We represent the shape using a spherical harmonic decomposition and then operate on the expansion coefficients to quantify the structural symmetry thanks to a mismatch function. Overall, we developed new mathematical and computational frameworks for symmetry detection using the polynomial expansion approach.

7.21. Pepsi-Dock: fast predictions of putative docking poses using accurate knowledge-based potentials functions to describe interactions between proteins

Participants: Emilie Neveu, Sergei Grudinin, David W. Ritchie, Petr Popov.

Many biological tasks involve finding proteins that can act as an inhibitor for a virus or a bacteria, for example. Such task requires knowledge on the structure of the complex to be formed. Protein Data Bank can help but only a small fraction of its proteins are complexes [16]. Therefore, computational docking predictions, being low-cost and easy to perform, are very attractive if they describe accurately the interactions between proteins while being fast to find which conformation will be the most probable. We have been developing a fast and accurate algorithm that combines the FFT-accelerated docking methods [67] with a precise knowledge-based potential functions [58] describing interactions between the atoms in the proteins .

Interactions between proteins follow complex and non-linear laws which computation is time-consuming. It is of common usage to start the predictions with a simple, approximated, expression of these interactions to then reduce the space search in order to use more complex laws. However we think it is important to use the most accurate free energy not to miss some important docking solutions. Thus, our aim is to integrate the very-detailed knowledge-based potentials into the *Hex* code and to take advantage of its exhaustive search, which is by now still the most efficient and reliable search algorithm [67] .

Last year, we adapted the machine learning process so that the knowledge-based potentials describing atom interactions can be translated into the polynomial basis used in *Hex*. The current evaluations of the knowledge-based scores takes more time than a shape+electrostatic representation but is still fast: exploring 10^9 conformations of a complex takes on average 5-10 minutes on a regular laptop computer.

This year, we run cross-validation experiments and tested different data sets in order to improve the predictions. Using bound conformations of each proteins to make the predictions, we retrieve up to 70% correct complexes of about 200 complexes. Results show that the knowledge-based potentials, while being general, correctly predict the interactions. Even better results could be achieved without the limitations in the search range by the spherical sampling grid which lacks of precision far away of its origin. Because many complexes have separation distances greater than 30 Å, we are now working on a multi-centre definition of the potentials in order to correctly predict the structures of protein complexes starting from their unbound structures.

7.22. Pepsi-Piper: rigid docking predictions using Pepsi potentials into Piper code

Participants: Sergei Grudinin, Emilie Neveu, Dima Kozakov, Dzmitry Podgorny.

This work is the continuation of the Pepsi-Dock project that aims to develop fast predictions of putative docking poses using accurate knowledge-based potentials functions to describe interactions between proteins. The goal is to integrate the precise, and yet easy to compute, distance-based pairwise knowledge-based potentials [58] into the Piper search code [48] in order to compare its exhaustive search with the *Hex* one. The former samples the conformations using a cartesian grid while the latter, a spherical one. We proved our potential used in *Hex* can predict the structures of complexes with a really good success rate, the main limitation being the lack of precision of the spherical sampling when the separation distance of the two proteins is too large. We think predicting docking combining our potential and a sampling search based on a cartesian grid as in Piper will achieve greater results, but will require more computational time.

We first adapt our potential to the Piper code and showed that the ranking results on the data set used for training are better than the ranking provided by Piper [25]: when the potential is used to sort the conformations, the correct solution is found in the first ten for 85% cases, while Piper found it in only 25% cases. The next step is to use the cartesian sampling to make docking predictions. When the Piper code will be ready to integrate our potential, we will be able to confront with other knowledge-based potentials such as the one initially used in Piper, DARS.

7.23. Flex-Dock: towards flexible docking predictions using metaheuristics optimisation methods

Participants: Emilie Neveu, Sergei Grudinin, Alexandre Hoffman, Angelo Migliosi, Xavier Besseron, Grégoire Danoy, Pascal Bouvry.

Docking numerical methods are used to predict the preferred location of one molecule with respect to the second when bound to each other. This is particularly useful for the design of drugs that inhibit the effects of viruses or bacteria. However proteins change their conformation upon binding and searching for flexible conformations involves enormous degrees of freedom and complex physics. Thus, the prediction of realistic interactions with full flexibility of the two partners is an intractable global optimisation problem.

There are currently several algorithms that produce high quality predictions of molecular complexes [43]. But very few manages to deal with the flexibility of the proteins. A common method is to refine the most probable predicted rigid complexes with a scoring allowing for flexibility [81]. Here, we want to tackle flexibility and sampling all together. Exhaustive search methods, which were by now the most accurate optimisation method for relatively small molecules [53] will be too time-consuming when it comes to large proteins. There is a strong need to explore and define new optimisation algorithms such as metaheuristic ones that can deal with several local minima and a large minima and a large search space. The main goal of this project is to define the problem and find for the optimisation method that will potentially give better results than the actual reference, SwarmDock [54].

We worked on a first comparison of several evolutionary-based algorithms (Genetic Algorithm [40], Differential Evolution [76], Particle Swarm Optimisation [46]) using rigid proteins only and on the use of multi-objective algorithms when the proteins are flexible.

To take into account flexibility, we approximate large-scale deformations of each proteins using an elastic network model combined with a low-frequency approximation called normal mode analysis such as in [81] or in [54]. Combined with the rigid transformation between the two proteins, it defines a complete while reduced set of degrees of freedom to search for.

The scoring function has to discriminate correct conformations from impossible ones. Our scoring is the main difference with SwarmDock. It takes into account the energy gained by docking using the precise knowledge-based potentials derived in [58], whereas only a simple physics-based energy is used in SwarmDock. We also want to explore another scoring that will also add the energetic cost of each moves of the proteins. To do so, we started to develop multi-objective algorithms. Combined with a Pareto Front analysis, this will help us to validate the scoring and to compare different evolutionary-based algorithms. Tests will be directly made on the Protein-Protein Benchmark [42] so that we can compare with other docking methods.

7.24. FastRMS: rapid determination of RMSDs corresponding to macromolecular rigid body motions, adding flexibility via collective motions

Participants: Sergei Grudin, Petr Popov, Emilie Neveu.

Computing the root mean sum of squared deviations (RMSDs) between two sets of coordinates each describing a different conformation of a macromolecule is a necessary step in many structural bioinformatics and molecular modelling technics to assess structural predictions [43], identify binding sites [49] or structurally classify proteins. A straightforward and universally-used method determines the RMSD with a computational complexity proportional to the number of atoms in the molecule. We recently presented RigidRMSD, a fast algorithm that determines RMSDs corresponding to a set of rigid body motions of a macromolecule in constant time with respect to the number of atoms in the molecule [57]. Here, we extend it to proteins with flexibility modelled with collective motion such as an elastic network model combined with normal mode analysis.

With these new assumptions, the complexity of the algorithm depends linearly or quadratically with the number of collective motion vectors selected to approximate the flexibility. The typical number of vectors needed to have accurate flexible movements being much lower than the number of atoms composing the molecules, we prove our algorithm is still faster than the common method. Our algorithm is particularly useful for rigid body modelling applications such as rigid body docking procedures allowing for flexibility via collective motions: clustering, high-throughput analysis and simulation results [49], [26], [59]. A C++ implementation of our algorithm will be soon available at <http://nano-d.inrialpes.fr/software/RigidRMSD>.

7.25. SAM : Spherical Polar Fourier Assembly of Protein Complexes with Arbitrary Point Group Symmetry

Participants: David W. Ritchie, Sergei Grudin.

We presented a novel FFT-based *ab initio* docking algorithm called “SAM” for building perfectly symmetrical models of protein complexes with arbitrary point group symmetry. The basic approach uses a novel and very fast 1D symmetry-constrained spherical polar Fourier search to assemble cyclic C_n systems from a given protein monomer. Structures with higher order (D_n , T , O , and I) point group symmetries may be built using a subsequent symmetry-constrained Fourier domain search to assemble trimeric sub-units. Our results show that the SAM algorithm can correctly assemble monomers of up to around 500 residues to produce a near-native complex structure with the given point group symmetry in 17 out of 18 test cases. The SAM program may be downloaded for academic use at <http://sam.loria.fr/>.

7.26. KSENIA : Knowledge of Native Protein-Protein Interfaces is Sufficient to Construct Predictive Models for the Selection of Binding Candidates

Participants: Petr Popov, Sergei Grudin.

Selection of putative binding poses is a challenging part of virtual screening for protein-protein interactions. Predictive models to filter out binding candidates with the highest binding affinities comprise scoring functions that assign a score to each binding pose. Existing scoring functions are typically deduced collecting statistical information about interfaces of native conformations of protein complexes along with interfaces of a large generated set of non-native conformations. However, the obtained scoring functions become biased toward the method used to generate the non-native conformations, i.e. they may not recognize near-native interfaces generated with a different method.

Present study demonstrates that knowledge of only native protein-protein interfaces is sufficient to construct well-discriminative predictive models for the selection of binding candidates. Here, we introduce a new scoring method that comprises a knowledge-based potential called *KSENIA* deduced from the structural information about the native interfaces of 844 crystallographic protein-protein complexes. We derive *KSENIA* using convex optimization with a training set composed of native protein complexes and their near-native conformations that are obtained using deformations along the low-frequency normal modes. As a result, our knowledge-based potential has only marginal bias toward a method to generate putative binding poses. Furthermore, *KSENIA* is smooth by construction, which allows to use it along with a rigid-body optimization to refine the binding poses. Using several test benchmarks we demonstrate that our method discriminates well native and near-native conformations of protein complexes from the non-native ones. Our methodology can be easily adapted to the recognition of other types of molecular interactions, such as protein-ligand, protein-RNA, etc. *KSENIA* will be made publicly available as a part of the SAMSON software platform at <https://www.samson-connect.net>.

7.27. Predicting Binding Poses and Affinities in the CSAR 2013–2014 Docking Exercises Using the Knowledge-Based Convex-PL Potential

Participants: Sergei Grudinin, Petr Popov, Emilie Neveu, Georgy Cheremovskiy.

The 2013–2014 CSAR docking exercise was the opportunity to assess the performance of the novel knowledge-based potential we are developing, named Convex-PL. The data used to derive the potential consists only of structural information from protein-ligand interfaces found in the PDBBind database. As expected, our potential proved to be very efficient in the near-native pose detection exercises, where we correctly predicted two near-native poses in the 2013 exercise and also ranked 22 near-native poses first and 2 second in the 2014 exercise. Somewhat more surprisingly, we obtained a fair performance in some of the CSAR affinity ranking exercises, where the Spearman correlation coefficients between our predictions and the experiments are greater than 0.5 for several protein–ligand sets. Nonetheless, affinity prediction exercises turned out to be a challenge, and significant progress in the development of our method is needed before we can successfully predict binding constants.

7.28. Prediction of homo- and hetero-protein complexes by ab-initio and template-based docking: a CASP-CAPRI experiment

Participants: Sergei Grudinin, Petr Popov, Emilie Neveu.

We present the results for CAPRI Round 30, the first joint CASP-CAPRI experiment, which brought together experts from the protein structure prediction and protein-protein docking communities. The Round comprised 25 targets from amongst those submitted for the CASP11 prediction experiment of 2014. The targets included mostly homodimers, a few homotetramers, and two heterodimers, and comprised protein chains that could readily be modeled using templates from the Protein Data Bank. On average 24 CAPRI groups and 7 CASP groups submitted docking predictions for each target, and 12 CAPRI groups per target participated in the CAPRI scoring experiment. In total more than 9500 models were assessed against the 3D structures of the corresponding target complexes. Results show that the prediction of homodimer assemblies by homology modeling techniques and docking calculations is quite successful for targets featuring large enough subunit interfaces to represent stable associations. Targets with ambiguous or inaccurate oligomeric state assignments, often featuring crystal contact-sized interfaces, represented a confounding factor. For those,

a much poorer prediction performance was achieved, while nonetheless often providing helpful clues on the correct oligomeric state of the protein. The prediction performance was very poor for genuine tetrameric targets, where the inaccuracy of the homology-built subunit models and the smaller pair-wise interfaces severely limited the ability to derive the correct assembly mode. Our analysis also shows that docking procedures offer a clear advantage over standard homology modeling techniques and that highly accurate models of the protein components are not always required to identify their association modes with acceptable accuracy.

Most of the targets in Round 30 of CAPRI were homodimers and homotetramers, thus it was a good opportunity to test our novel symmetry assembling docking method. To do so, we imposed C2 symmetry constraints for all the homodimers and we imposed C4 and D2 symmetry constraints for all the homotetramers from the target complexes. Below, we present the new fast multi-resolution method for docking both symmetric and non-symmetric protein complexes that was used in Round 30 of CAPRI. First, the structures of the individual subunits were taken from the stage two predictions of the CASP10 assessment experiment. More precisely, starting from 150 available CASP 3D models of monomers, we predicted models of symmetric multimers using the novel symmetry docking method, which performs symmetry-induced protein docking using the shape-complementarity scoring function computed as spherical polar Fourier correlations. Specifically, this method performs exhaustive search over the available (four in case of cyclic symmetries or six otherwise) degrees of freedom for the given point group symmetry type. For the targets of Round 30 of CAPRI we imposed three types of symmetry, C2, C4, and D2. For the case of heterodimers, we used the standard Hex docking method.

For the input of the docking methods, we generated the scaffolds of initial models of monomers by cutting-off the side chains. More specifically, we mutated all side-chains except for the glycines to alanines. Compared to the standard all-atom rigid-body docking methods, we expect the scaffold docking approach to produce binding poses that are less sensitive to the flexibility of the side-chains. We clustered the solutions with the threshold ligand-RMSD value of 8 Å using the RigidRMSD library. Finally, we ranked the clusters by the value of the best score and kept 50 best clusters for the refinement stage. In total, for each target we proceeded to the refinement with 7,500 modeled structures of protein complexes.

On the next step, we optimized each putative binding interface of the all-atom representation of a protein complex by means of a rigid-body first-order minimization scheme. Specifically, after each rigid-body minimization step we proceeded with the optimization of side-chains described by the rotameric representation using the SCWRL4 package. We computed the interactions between the subunits in a protein complex using the novel reference state-free knowledge-based scoring function KSENIA, which is smooth by construction and is thus very suitable for a gradient-based minimization protocol. Finally, we ranked the predictions by the value of the KSENIA potential of the optimized structure and selected ten best candidates for the submission.

7.29. Convex relaxation for non-convex quadratic optimization problems with applications to side-chain prediction in protein structures

Participants: Aleksandr Katrutsa, Sergei Grudinin.

The side-chain prediction problem is the major part of the more general protein structure prediction problem, which is very important for drug design and in the prediction of stable protein mutations. Formally, the side-chain prediction problem states in the form of discrete quadratic optimization problem with an indefinite matrix in the quadratic term,

$$\mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{b}^\top \mathbf{x} \rightarrow \min_{\mathbf{x} \in \{0,1\}^n} \quad (8)$$

This problem is NP-hard, so to get a good approximation solution we used convex semidefinite relaxation with different types of constraints. This approach is the powerful optimization technique that helps to reformulate the initial non-convex problem as a convex one and sometimes even gives the exact solution. The important step is to operate with precise energy function, which is used to compute the energy of different interactions

in proteins. To obtain this, we used the machine-learning procedure, which extracts the parameter vector for the potential from the training set of protein structures. After the training step, we used this vector to compute the energy of a protein and to find the side-chains corresponding to the minimal total energy of the protein. The current accuracy in side-chain prediction is about 80%, which is achieved using the spectrum relaxation of the matrix in the quadratic term. Also, this approach is very fast, precisely, it requires less than 1 second per protein to predict the positions of its side-chains.

7.30. Critical assessment of protein-ligand docking methods via Drug Design Data Resource (D3R) Challenge 2015

Participants: Andreas Eisenbarth, Sergei Grudinin.

We participated at the Drug Design Data Resource (D3R) Challenge 2015. In the challenge, we were given protein structures and sets of ligand molecules in order to detect the putative binding poses. The aim was to find the energetically most favourable pose of each ligand relative to a protein. To do so, we first performed the docking simulations using the state-of-the-art software AutoDock Vina, then explored sets of parameters that produced chemically reasonable poses, and finally did the re-scoring using the ConvexPL potential. Later, we critically examined AutoDock Vina sampling method and detected points where it can be improved and also assessed the integration of our inhouse developed ConvexPL scoring algorithm.

7.31. Towards the development of FFT-accelerated flexible fitting methods

Participants: Alexandre Hoffmann, Valerie Perrier, Sergei Grudinin.

We studied a set of new methods for non-rigid molecular fitting. The problem can be formulated as follows : Let \mathcal{P}_1 and \mathcal{P}_2 be two molecular structures (e.g. proteins). We are given $d_1 : \mathbb{R}^3 \mapsto \mathbb{R}$, the electron density of \mathcal{P}_1 and $(Y_k \in \mathbb{R}^3)_{k=1 \dots N_{atoms}}$, the average positions of the atoms of \mathcal{P}_2 . Assuming we can generate an artificial electron density $d_2 : \mathbb{R}^3 \mapsto \mathbb{R}$ from $(Y_k \in \mathbb{R}^3)_{k=1 \dots N_{atoms}}$, our problem is to find a transformation of the atoms $T : \mathbb{R}^{3N_{atoms}} \mapsto \mathbb{R}^{3N_{atoms}}$ that minimizes the L^2 distance between d_1 and d_2 .

In image processing, this problem is usually solved using the optimal transport theory, but this method assumes that both densities have the same L^2 norm, which is not necessarily the case for the fitting problem. To solve this problem, one instead starts by splitting T into a rigid transformation T_{rigid} (which is a combination of translation and rotation) and a flexible transformation T_{flex} . Two classes of methods have been developed to find T_{rigid} :

- the first one uses optimization techniques such as gradient descent,
- the second one uses the Fast Fourier Transform (FFT) to compute the Cross Correlation Function (CCF) of d_1 and d_2 .

We have developed several algorithms based on the FFT to find T_{rigid} and we have developed two algorithms for flexible molecular fitting that are based on convex and non-convex optimization and the trust region methods. Our tests demonstrate that while one method gives good results for small deformations, the other gives good results for bigger deformations.

We have been also improving the current NMA method (which is essentially a model reduction technique), that is used in other tools such as the flexible fitting to small angle scattering profiles. Finally, we started the development of a method for a harder fitting/docking problem in which only electron density would be known. The basic idea would be to find the C^1 -diffeomorphism $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ that minimizes the L^2 distance between d_1 and d_2 .

We developed several stand-alone C++ libraries to solve some of our problems including:

- a non-convex optimization library,
- a normal mode analysis library,
- a fitting library that implements our new methods.

NECS Project-Team

7. New Results

7.1. Network systems and graph analysis

7.1.1. Distributed estimation of graph Laplacian eigenvalues

Participants: A. Kibangou [Contact person], T.-M. D. Tran.

Linear average-consensus is a well-known iterative protocol allowing agents to converge to the average of initial values by taking suitable convex combinations of the messages received from neighbors. From the recent literature, it is known that, after a finite time, some consecutive measurements of a state of the consensus dynamical system can be used to compute the exact average of the initial condition. In [23], we have shown that these measurements can also be used for estimating the Laplacian eigenvalues of the graph representing the network. As recently shown in the literature, by solving the factorization of the averaging matrix, the Laplacian eigenvalues can be inferred. In our paper, the problem is posed as a constrained consensus problem. A first formulation (direct approach) yields a non-convex optimization problem, which we solve in a distributed way using Lagrange multipliers. A second formulation (indirect approach) is obtained after a suitable reparameterization. The problem is then convex and is solved by using the distributed subgradient algorithm and the alternating direction method of multipliers (ADMM). The proposed algorithms allow estimating the actual Laplacian eigenvalues with high accuracy. However, they face numerical instability when considering very large graphs.

7.1.2. Distributed solution to the network reconstruction problem

Participants: A. Kibangou [Contact person], T.-M. D. Tran.

We address the problem of reconstructing the network topology from data propagated through the network by means of a linear average-consensus protocol. In [34], we propose a new method based on the distributed estimation of graph Laplacian spectral properties. Precisely, the identification of the network topology is implemented by estimating both eigenvalues and eigenvectors of the consensus matrix, which is related to the graph Laplacian matrix. Having already solved in [23] the problem of estimating the eigenvalues (see paragraph above), in this paper we focus on the eigenvectors. We show how the topology can be reconstructed in presence of anonymous nodes, i.e., nodes that do not disclose their ID. Actually, in presence of anonymous nodes, eigenvectors are estimated up to a permutation of rows; the obtained graph is then isomorphic to the original one. Moreover, under some observability assumption on the consensus dynamical system (if the graph is node-observable or neighborhood-observable from the node of interest) and if all the entries of the initial condition of the network state are distinct, then the node can exactly reconstruct the network topology. If the entries of the initial condition of the network state are independently generated from a continuous probability distribution, then the node can reconstruct the network topology almost surely. The main assumption in this work is: all eigenvalues are distinct, that is the case of most random graphs. Future works encompass the design of the network reconstruction protocol that deals with spectrums in which the multiplicities of the eigenvalues can be higher than 1 and also directed graphs. In addition, numerical issues for large graphs are to be considered for making the proposed method scalable.

7.2. Sensor networks: estimation and data fusion

7.2.1. Multisensor data fusion for attitude estimation

Participants: H. Fourati [Contact person], A. Kibangou, A. Makni, T. Michel, P. Geneves [Tyrex, Inria], N. Layaida [Tyrex, Inria].

Multisensor data fusion has gained in importance over the last decades and found applications in an impressive variety of areas within diverse disciplines: navigation, sensor networks, intelligent transportation systems, security, medical diagnosis, biometrics, environmental monitoring, remote sensing, measurements, robotics, and so forth. Different concepts, techniques, and architectures have been developed to optimize the overall system output in applications for which sensor fusion might be useful and enables development of concrete solutions. These concepts and ideas are treated in the book [35], as a response to the great interest and strong activities in the field of multisensor data fusion during the last few years, both in theoretical and practical aspects.

In the team, we have carried out works related to attitude estimation for pedestrian navigation purpose.

In [32], we investigated a new modeling and filtering approach for rigid body attitude estimation. In contrast to the current state-of-the-art. where the process model is driven by gyroscope measurements, we propose an alternative modeling formulation where the process model is fed by the magnetometer measurements. The resulting dynamic model takes the form of a descriptor system, also known as singular system. Based on this model and using the quaternion formalism we derive a recursive filter whose performance is validated through numerical and experimental tests.

In [20], we focused on two main challenges. The first one concerns the attitude estimation during dynamic cases, in which external acceleration occurs. In order to compensate for such external acceleration, we design a quaternion-based adaptive Kalman filter q-AKF. Precisely, a smart detector is designed to decide whether the body is in static or dynamic case. Then, the covariance matrix of the external acceleration is estimated to tune the filter gain. The second challenge is related to the energy consumption issue of gyroscope. In order to ensure a longer battery life for the Inertial Measurement Units, we study the way to reduce the gyro measurements acquisition by switching on/off the sensor while maintaining an acceptable attitude estimation. The switching policy is based on the designed detector. The efficiency of the proposed scheme is evaluated by means of numerical simulations and experimental tests.

In [33], we investigated the precision of attitude estimation solutions in the context of Pedestrian Dead-Reckoning (PDR) with commodity smartphones and inertial/magnetic sensors by carrying out a concise comparison of various methods. We conducted an experimental study with a precise ground truth obtained with a motion capture system. We precisely quantified the error in attitude estimation obtained with each filter which combines a 3-axis accelerometer, a 3-axis magnetometer and a 3-axis gyroscope measurements.

7.2.2. Sensor placement of unreliable sensors

Participants: F. Garin [Contact person], P. Frasca [U. Twente], B. Gerencsér [U. Catholique de Louvain], J. Hendrickx [U. Catholique de Louvain].

We consider problems in which sensors have to be deployed in a given environment in such a way to provide good coverage of it. It is clear that sensor failures may deteriorate the performance of the resulting sensor network. Then, it is also natural to ask if taking into account such uncertainties changes the coverage optimization problem and leads to a different optimal solution. For simplicity, we start considering a one-dimensional problem, where sensors are to be placed on a line in such a way to optimize the disk-coverage cost. The optimal solution for reliable sensors is simply an equally-spaced configuration of the sensors. If we allow that the sensors may fail to take or communicate their measurements, this solution may instead not be optimal. In our work, we assume that sensor can fail, independently and with a same failure probability, and we aim to minimize, in expectation, the largest distance between a point in the environment and an active sensor. Our first result states that the problem at hand is equivalent to a linear program, albeit with a number of variables growing exponentially with the number of sensors. This fact allows for a computational solution that is tractable if the number of sensors is not large. Secondly, we show that for large number of sensors n , the cost of the equispaced placement is asymptotically optimal, i.e., the ratio between its cost and the optimal cost tends to 1 when n grows. By contrast, we show in that a random sensor placement has an expected cost which is larger. This work is described in the paper [18].

7.3. Control design and networked control

7.3.1. Control design for hydro-electric power-plants

Participants: C. Canudas de Wit [Contact person], S. Gerwig, F. Garin, B. Sari [Alstom].

We have a collaboration with Alstom on collaborative and resilient control of hydro-electric power-plants, with the CIFRE PhD thesis of Simon Gerwig. The first goal of this research is to improve performance of a hydro-electric power-plant outside its design operation conditions, by cancellation of oscillations that occur in such an operation range. Indeed, current operation of power-plants often requires to operate on a variety of conditions, often different from the ones initially considered when designing the plant. At off-design operation pressure, the hydraulic turbine exhibits a vortex rope below the runner. This vortex generates pressure fluctuations after the turbine and can excite the hydraulic pipes. Indeed the water is compressible and the pipe walls elastic, so the system can oscillate. The goal is to damp these pressure oscillations as they create vibrations in the system and can lead to damages. Our first contribution has been to model the effect of the vortex rope on the hydraulic system as an external perturbation source acting on pipes. The pipes themselves are described with equations taking into account water compressibility and pipe-wall elasticity. The resulting model is nonlinear with hyperbolic functions in the equations (analogous to high-frequency transmission lines), from which we obtain a suitably linearized model.

7.3.2. Collaborative source seeking

Participants: C. Canudas de Wit [Contact person], R. Fabbiano, F. Garin.

The problem of source localization consists in finding, with one or several agents possibly cooperating with each other, the point or the spatial region from which a quantity of interest is being emitted. Source-seeking agents can be fixed sensors, that collect and exchange some information about the signal field and try to identify the position of the source (or the smallest region in which it is included), or moving devices equipped with one or more sensors, that physically reach the source in an individual or cooperative way. This is particularly difficult when the agents have limited or no position information and GPS navigation is not available, as in underwater navigation or in cave exploration: for instance, source localization is relevant to many applications of vapor emitting sources such as explosive detection, drug detection, sensing leakage or hazardous chemicals, pollution sensing and environmental studies. Other fields of interest are sound source localization, heat source localization and vent sources in underwater field. Techniques present in literature either are based on a specific knowledge of the solution of the diffusion process, or make use of an extremum-seeking approach, exciting the system with a periodic signal so as to explore the field and collect enough information to reconstruct the gradient of the quantity of interest. Our approach lies in the computation of derivatives (potentially of any order) from Poisson integrals that, for isotropic diffusive source in steady-state, whose solution satisfies the Laplace equation, allows for a gradient search with a small computation load (derivatives are computed by integrals) and without requiring any knowledge of the closed-form solution, avoiding in the same time extremum-seeking oscillations; this has the additional advantage of an intrinsic high-frequency filtering, that makes the method robust to measurement noise. We also propose a distributed version of this algorithm, where agents communicate in order to reconstruct gradient information from local pointwise measurements, and a control law combines the two objectives of formation control (to have a circular formation, so that measurements are taken around circle) and gradient ascent (so as to move towards the source); differently from previous literature, the moving agents do not need to know their absolute position, but only relative bearing angle of their neighbours. This work is the topic of the Ph.D. thesis of Ruggero Fabbiano [12].

7.3.3. Synchronization of heterogenous networks

Participants: E. Lovisari [Contact person], C.-Y. Kao [National Sun Yat-Sen University, Taiwan].

Synchronization of agents in large-scale networks is studied in [19]. Each agent is modeled as a Single Input Single Output operator composed of the series of a common Linear Time-Invariant system and a possibly nonlinear perturbation. Interconnection is represented via a graph whose edges model communication channels between agents, in turn modeled as a nominal component and a possibly nonlinear perturbation. Two agents are synchronized if their outputs are the same, possibly time-varying signal. The main result provides synchronization certificates based on the Robust Control Technique of Integral Quadratic Constraints. Exploitation of graph structures allows then to reduce the computational burden of the certificate in a way that scales with the dimension of the network. This provides framework which unifies and extends several results already presented in the literature.

7.3.4. *Observer-based FDI scheme for switched systems with sensor faults*

Participants: H. Fourati [Contact person], D. E. C. Belkhiat [U. Setif], D. Jabri [U. Setif].

The Fault Detection and Isolation (FDI) problem for a class of Switched Linear Systems (SLS) subject to sensor faults and unknown bounded Disturbances is proposed in [24]. The main work is based on the design of a generalized switched observer scheme. The FDI problems have been solved by using a robust control techniques. A suitable trade-off between the robustness to disturbances and the sensitivity to sensor faults was obtained. The main results are reformulated by using Linear Matrix Inequality (LMI) formulation. An example is included to illustrate the efficiency of the proposed approach.

7.4. Transportation networks and vehicular systems

7.4.1. *Traffic estimation: sensors placement and data fusion*

Participants: C. Canudas de Wit [Contact person], E. Lovisari, A. Kibangou.

Ability to reconstruct the state of a transportation network is of paramount importance. Indeed, such an information is used to forecast traffic evolution, to inform drivers in real-time through navigation systems, to provide statistical information to public authorities to detect in a timely fashion accidents and predict hazardous scenarios, and finally to compute controls and to actuate the network through traffic lights, ramp metering, or adaptive speed limits.

A primary source of information on the state of the network are fixed traffic detectors, namely, devices able to measure density, flow and average speed of vehicles crossing the section of the road where they are placed. We have addressed the Optimal Sensor Placement problem [31], namely, the problem of finding the best physical location for sensors. This is based on a trade-off of two contrasting objectives: the first, to maximize the performance of state reconstruction; the second, to minimize the total economic cost of the network. To simplify the setting, we consider the related problem of reconstruction in a static setting, by considering as performance metric the error covariance of an estimator of the cumulative flows in the network over a long period of time. Since the resulting trade-off problem remains a combinatorial problem, we relax it using a method that we call Virtual Variance algorithm, based on the idea to associate to each sensor a virtual variance, which is large when the sensor is not needed for good reconstruction of the flow vector. The only input that the algorithm needs is an estimate of the matrix of splitting ratios and the nominal variance of each sensor. Since in real application a pre-existing sensor network is often unavailable, possible alternatives are field surveys with operators visually counting vehicles, as commonly done for calibration of traffic software, or temporary non-invasive equipment such as radar traffic detectors.

In addition to fixed traffic detectors, the spread of wireless devices allows new sensing and communication capabilities. In particular, for the traffic application, any vehicle equipped with a GPS device can act as a probe in the traffic and provide Floating Car Data (FCD). If a non negligible fraction of vehicles acts as probe, the collected data provides an estimate of the evolution of speed in the network. Due to privacy reasons, single vehicles traces are usually not directly used, but rather aggregated as average speed of vehicles in segments of road. Advanced methodologies, such as the one used by INRIX, ensure a very fine spatial partition of the network, with segments as short as 250 meters (see the INRIX official website <http://www.inrix.com/xd-traffic>). Compared to fixed sensors, this technology is less precise, but since it exploits existing communication

systems it is relatively less expensive and already covers all major traffic networks. In our work [30], [29], we propose an algorithm that aims at reconstructing the traffic density by fusing fixed sensors measurements and Floating Car Data. We employ a macroscopic model, partitioning the network in cells and assigning to each cell a density of vehicles. The latter evolves dynamically according to a first order mass-conservation law. Our approach inherits from the CTM the cell-based topology, but we do not directly employ the resulting dynamical model. Instead, inflows and outflows are estimated on the basis of the available flow measurements only, and speed measurements are employed to compute a pseudo-measurement of the density. These quantities are the inputs for the density observer. In addition, we propose a gradient descent method to calibrate the Fundamental Diagram, and we implement the proposed solution using real fixed sensor measurements from the Grenoble Traffic Lab [14] and speed FCD measurements provided by INRIX, one of the most well known traffic solutions companies.

7.4.2. Traffic forecasting

Participants: A. Kibangou [Contact person], C. Canudas de Wit, H. Fourati, A. Ladino Lopez.

Traffic forecasting is one of the most desired tools for traffic management, requested by operators and commuters. In the era of data deluge in which we are, measurements collected by sensors are important sources of information that require analysis, classification and processing in order to detect patterns and behaviours that can be exploited for traffic prediction ([30], [37]). The collected information can be classified by clustering algorithms such as K-means; each cluster collects traffic patterns, which in some cases characterize typical regimes such as congestion. Based on clustered data, we have first developed forecasting schemes based on adaptive Kalman filtering [14]. These schemes were designed for specific origin-destination (OD) pairs, assuming availability of measurements whatever the time instants. Recently, within the PhD thesis in progress of Andres Ladino Lopez, we considered a network-oriented forecasting scheme, where travel time measurements are assumed to be available only for a few sets of OD pairs and sporadically (missing data), but forecasting is to be achieved for all the OD pairs of the network. To reduce the dimensionality of the problem, we actually predicted the travel time for the internal state of the network. In addition, since travel time measurements for all the OD pairs cannot be available all the time, we faced a missing data problem. To overcome this issue, we resorted to a data imputation based on a dictionary learning approach. From the imputed data, a clusterization was achieved, defining different clusters characterized by a centroid containing the mean of the data and a given dispersion around it. The evolution of the centroid can be used as future observation, herein called pseudo-observation, that can feed a Kalman filter. Therefore the prediction problem was solved as a filtering one. However, the main question was, how to associate the current day data to a specific cluster, since we didn't know its future? To solve this issue, we run Kalman filters for each cluster and then made the fusion of the obtained forecasts.

7.4.3. Traffic control

Participants: C. Canudas de Wit [Contact person], F. Garin, D. Pisarski, P. Grandinetti, E. Lovisari, G. Como [U. Lund], K. Savla [U. of Southern California].

The activities of the team on traffic control can be organized in three parts: freeway traffic control, urban control, and analysis and control of monotone flows.

First, we have studied optimal balancing of vehicle density in the freeway traffic. The optimization is performed in a distributed manner by utilizing the controllability properties of the freeway network represented by the Cell Transmission Model. By using these properties, we identify the subsystems to be controlled by local ramp meters. The optimization problem is then formulated as a non-cooperative Nash game that is solved by decomposing it into a set of two-players hierarchical and competitive games. The process of optimization employs the communication channels matching the switching structure of system interconnectivity. By defining the internal model for the boundary flows, local optimal control problems are efficiently solved by utilizing the method of Linear Quadratic Regulator. The developed control strategy is tested via numerical simulations in two scenarios for uniformly congested and transient traffic. This work is described in the paper [21].

Second, we have considered optimal or near-optimal operation of traffic lights in an urban area. The goal is on-line optimization of traffic light schedule in real time, so as to take into account variable traffic demands, with the objective of obtaining a better use of the road infrastructure. More precisely, we aim at maximizing total travel distance within the network, while also ensuring good servicing of demands of incoming cars in the network from other areas. One way to address the complexity of the resulting optimization problem is to use a simplified averaged model for the traffic variables, and to optimize only the duty-cycles of traffic lights, i.e., the fractions of green time. This, together with a one-step optimization horizon, allows us to turn the problem into a simple linear program [27]. Another approach is to include as optimization variables both duty-cycles and phases of the traffic lights. We show how to turn the resulting problem into a mixed-integer linear program (MILP). Then, to overcome its complexity, we propose a sub-optimal distributed solution, while the global MILP can be used off-line for performance comparison [28].

Third, stability and throughput properties of monotone dynamical flow networks are studied in [15]. Vehicular density on the cells of the networks evolves according to laws that deterministically split the traffic flow at each intersection as a function of the density of other cells around the intersection. By exploiting the theory of monotone operators it is proven that under certain condition the system achieves an equilibrium that maximizes the throughput of the network, namely, if the inflow is smaller than the network capacity, then asymptotically the total outflow matches the total inflow, otherwise the total outflow matches the network capacity. In [25] a different traffic model is employed which uses demand and supply functions to relate density and flows of the network. The Social Optimum Dynamic Traffic Assignment, which is an optimal control problem with cost corresponding to the total travel time of vehicles in the network, is solved making use of ramp metering and speed limits. The optimization is shown to be a convex optimization problem, making its numerical solution feasible by employing well known tools.

7.4.4. Energy-aware control of communicating vehicles

Participants: C. Canudas de Wit [Contact person], G. de Nunzio.

The research in this domain focuses mainly on efficient traffic energy consumption and has been carried out at two levels. Strategies for both the vehicles-side and the infrastructure-side eco-management have been proposed or extended. As for the vehicle-side control of communicating vehicles, assuming I2V communication, and therefore full knowledge of the traffic lights timings, the goal is to analyze the driving horizon and compute an energy-efficient speed advisory for the driver. As in previous works, stops at a red traffic light are to be avoided. The novelty of our approach is summarized as follows. Given a set of green traffic light phases, there exist different driving profiles to reach a given destination at a given final time in compliance with traffic lights constraints (i.e. always catching the green light) and city speed limits. The presented strategy is capable of an a priori identification of the most energy-efficient velocity trajectory, by approximating the available paths and their energy cost with an oriented weighted graph. The computational complexity of the graph creation has been reduced in this work from exponential [26] to polynomial, thanks to the introduction of the line graph. The computation time has been consequently significantly reduced. Only after this preliminary stage of path selection, a formal optimization problem is solved in order to calculate the optimal arrival times at each intersection, by explicitly minimizing the energy consumption of the vehicle. This approach qualifies as a pre-trip eco-driving ADAS, since the speed advisory is provided to the driver at the beginning of the driver horizon. However, given the very little computation time required by the algorithm, it may be employed online thus enabling in-trip assistance features. This allows to respond dynamically to traffic perturbations and/or deviations from the speed advisory, and to increase the robustness and the applicability of the strategy in a realistic environment. Simulations in a microscopic traffic simulator demonstrate that the proposed strategy is able to deal online with perturbations coming from traffic and to reduce the overall energy consumption without affecting travel time [16].

At a lower level, the eco-driving from the vehicle perspective has been also addressed in a comprehensive analysis of the optimal driving strategy for different types of powertrains [22].

As for the infrastructure-side eco-management, this year's research focused on extending the results published in [26]. The two-way arterial bandwidth maximization problem is addressed with a particular focus on the

benefits induced by the speed advisory, and on reducing energy consumption. The problem with internal offsets constraints presents difficulties that make necessary the formulation of the problem as an MILP. The first contribution of our work lies in the addition of terms representing traffic energy consumption and network travel time to the objective function of the two-way arterial bandwidth maximization. The segment speeds, as additional control action, allow to reach higher theoretical bandwidths but might induce driving discomfort and higher energy consumption if the variability of the recommended speeds is too high. Furthermore, optimal solutions with low speeds and high travel time are to be avoided, in trade-off with the energy consumption. The second contribution is given by the extensive evaluation of the benefits of bandwidth maximization via a microscopic traffic simulator. Bandwidth is a theoretical quantity and a correlation with known traffic performance metrics needs to be established in order to justify its use. The combined control of offsets and speed advisory is shown to have a large impact on energy consumption without affecting the travel time. Lastly, an analysis of the traffic performance at different levels of traffic demands has been conducted, testing both under-saturated traffic conditions with the existence of a green wave, and saturated conditions. The goal of this analysis is to identify the best operation conditions of the presented approach, assess the performance degradation with traffic load, and, most importantly, propose a demand-dependent optimization. Several strategies were compared to the presented one in order to assess its performance. This work has been submitted for review to the IEEE Transactions on Control Systems Technology.

Finally, a detailed description of the proposed strategies and the achieved results in the domain of the energy-aware traffic management in urban networks can be found in the dissertation [11].

NON-A Project-Team

6. New Results

6.1. Homogeneity Theory

Homogeneity is one of the tools we develop for finite-time convergence analysis. In 2015 this concept has received various improvements:

- The concept of homogeneous evolution equation in a Banach space has been introduced in [67]. It provides the background for the extension of all homogeneity-based tools for control design and analysis to distributed parameters systems.
- Scalability is a property describing the change of the trajectory of a dynamical system under a scaling of the input stimulus and of the initial conditions. Particular cases of scalability include the scale invariance and fold change detection (when the scaling of the input does not influence the system output). In the paper [19] is shown that homogeneous systems have this scalability property while locally homogeneous systems approximately possess this property.
- In the paper [25] the notion of homogeneity in the bi-limit is extended to local homogeneity and then to homogeneity in the multi-limit. The converse Lyapunov/Chetaev theorems on (homogeneous) system instability are obtained. The problem of oscillation detection for nonlinear systems is addressed. The sufficient conditions of oscillation existence for systems homogeneous in the multi-limit are formulated.
- The notion of weighted homogeneity is extended in [81] to the time-delay systems. It is shown that the stability/instability of homogeneous functional systems on a sphere implies the global stability/instability of the system. The notion of local homogeneity is introduced, a relation between stability/instability of the locally approximating dynamics and the original time-delay system is established using Lyapunov-Razumikhin approach
- In [27] global delay independent stability is analyzed for nonlinear time-delay systems by applying homogeneity theory. It is shown that finite-time stability can be encountered in this class of systems under uniformity of the convergence time with respect to delay. Some additional tools for stability analysis of time-delay systems using homogeneity are also presented: in particular, it is shown that if a time-delay system is homogeneous with nonzero degree and it is globally asymptotically stable for some delay, then this property is preserved for any delay value, which is known as the independent of delay (IOD) stability.
- Theorems on Implicit Lyapunov Functions for finite-time and fixed-time stability analysis of nonlinear systems are presented in [37]. Based on these results, new homogeneous nonlinear control laws are designed for robust stabilization of a chain of integrators. The presented results are extended to Multi-Input Multi-Output in [38]. A time-suboptimal control design algorithm based Implicit Lyapunov Function Method is developed in [40]. A robustness-oriented comparison of the optimal and suboptimal solutions in practical implementations of the proposed controller is performed via the numerical example of double integrator. A novel scheme of practical implementation of the implicit Lyapunov function-based control is developed in [79]. It replaces the implicitly defined Lyapunov function (in the feedback law) with the homogeneous norm of the state. Such a modification simplifies the practical application of the finite-time stabilizing feedback control.
- The uniform stability notion for a class of nonlinear time-varying systems is studied in [42] using the homogeneity framework. It is assumed that the system is weighted homogeneous considering the time variable as a constant parameter, then several conditions of uniform stability for such a class of systems are formulated. The results are applied to the problem of adaptive estimation for a linear system. The detailed report on time-varying homogeneity is given in [83].

- In the paper [52] we consider the continuous homogeneous observer defined in the case of the triple integrator. Originally, convergence of the algorithm was only proved when the degree of homogeneity was sufficiently close to 0 without more tractable information. We show here that, in the case of the triple integrator, the observer presents global finite-time stability for any negative degree under constructive conditions on the gains. This is achieved with a homogeneous Lyapunov function design.
- The work [61] addresses the stabilization of dynamical systems in presence of uncertain bounded perturbations using theory. Under some assumptions, the problem is reduced to the stabilization of a chain of integrators subject to a perturbation and is treated in two steps. The evaluation of the disturbance and its compensation. Homogeneous observer and control are the tools utilized to achieve a global asymptotic stability and robustness. The result is formally proven and, to validate the theory, it is applied to the control of the telescopic link of a hydraulic actuated industrial crane used in forestry.
- A geometric homogeneity of evolution equation in a Banach space is introduced in [67]. Scalability property of solutions of homogeneous evolution equations is proven. Some qualitative characteristics of stability of trivial solution are also provided. In particular, finite-time stability of homogeneous evolution equations is studied. Theoretical results are supported by examples from mathematical physics.
- The second order planar nonlinear affine control problem is studied [69]. A homogeneous robust finite-time stabilizing control is developed for the most general case of matched and, the more challenging, mismatched nonlinear perturbations. A homogeneous observer is designed for the planar system. Explicit restrictions on the observer gains and nonlinearities are presented. The main contribution lies in the proposed combination of the explicit and implicit Lyapunov function methods as well as weighted homogeneity while providing finite-time stability analysis.

6.2. Algebraic Technique For Estimation, Differentiation And Its Applications

Algebraic technique is the other tool we develop for providing finite-time convergence.

- The integer order differentiation by integration method based on the Jacobi orthogonal polynomials for noisy signals was originally introduced by Mboup, Join and Fliess. The paper [35] proposes an extension of this method from the integer order to the fractional order to estimate the fractional order derivatives of noisy signals. Two fractional order differentiators are deduced from the Jacobi orthogonal polynomial filter, using the Riemann-Liouville and the Caputo fractional order derivative definitions respectively. Exact and simple formulas for these differentiators are given by integral expressions. Some error bounds are provided for the corresponding estimation errors. The noise error contribution due to a large class of stochastic processes is studied in discrete case.
- Armed with structures, group sparsity can be exploited to improve the performance of adaptive estimation. In the paper [45], the adaptive estimation algorithm for cluster structured sparse signals, called A-CluSS, is proposed. In particular, a hierarchical Bayesian model is built, where both sparse prior and cluster structured prior are exploited simultaneously. The adaptive updating formulas for statistical variables are obtained via the variational Bayesian inference and the resulted algorithms can adaptively estimate the cluster structured sparse signals without knowledge of block size, block numbers and block locations. In [75], a group sparse regularized least-mean-square (LMS) algorithm is proposed to cope with the identification problems for multiple/multi-channel systems. An iterative online algorithm is proposed via proximal splitting method.

6.3. Set-Theoretic Methods of Control And Estimation

Interval and ellipsoidal estimations can be regarded as particular finite-time algorithms, since they provide guaranteed estimates of the values from the initial time. We develop these tools for some years now.

- An approach to interval observer design for Linear Parameter-Varying (LPV) systems is proposed in [20]. Stability conditions are expressed in terms of matrix inequalities. Applying L1/L2 framework the robustness and estimation accuracy with respect to model uncertainty are analyzed.
- New delay-dependent conditions of positivity for linear systems with time-varying delays are introduced in [56]. These conditions are applied to interval observer design for systems with time-varying delays in the state equations and in the measurements. In [28] the problem of interval observer design is addressed for a class of descriptor linear systems with time delays. An interval observation for any input in the system is provided. The control input is designed together with the observer gains in order to guarantee interval estimation and stabilization simultaneously. Efficiency of the proposed approach is illustrated by numerical experiments with Leontief delayed model.
- The work [29] is devoted to interval observers design for discrete-time Linear Parameter-Varying (LPV) systems under the assumption that the vector of scheduling parameters is not available for measurements. Two problems are considered: a pure estimation problem and an output stabilizing feedback design problem where the stability conditions are expressed in terms of Linear Matrix Inequalities (LMIs).
- The paper [48] investigates the interval observer design for a class of nonlinear continuous systems, which can be represented as a superposition of a uniformly observable nominal subsystem with a Lipschitz nonlinear perturbation. It is shown in this case there exists an interval observer for the system that estimates the set of admissible values for the state consistent with the output measurements. In [77] similar methodology is extended to singular systems.
- A finite-time version, based on Implicit Lyapunov Functions, for the Attractive Ellipsoid Method is developed in [65]. Based on this, a robust control scheme [36] is presented to ensure finite-time convergence of the solutions of a chain of integrators with bounded output perturbations to a minimal ellipsoidal set. The control parameters are obtained by solving a minimization problem of the " size " of the ellipsoid subject to a set of Linear Matrix Inequalities, and by applying the implicit function theorem.
- In [78] we consider a problem of sliding mode control design for LTI systems with multiplicative disturbances of the input and noisy measurements of the output. We apply the minimax observer to provide the best possible estimate of the system's state. Then we solve a problem of optimal reaching for the observer: we design sub-optimal control algorithms generating continuous and discontinuous feedback controls that steer the observer as close as possible to a given sliding hyperplane in a finite time.

6.4. Observability And Observer Design For Nonlinear Systems

- In [18] a method to carried out the state estimation is proposed for a class of nonlinear systems with unknown inputs whose dynamics is governed by differential-algebraic equations (DAE). We achieve, under suitable conditions, to replace the original DAE for a system with differential equations only by using a zeroing manifold algorithm inducing a state space dimension reduction.
- In the paper [44], we investigate the estimation problem for a class of partially observable nonlinear systems. For the proposed Partial Observer Normal Form (PONF), necessary and sufficient conditions are deduced to guarantee the existence of a change of coordinates which can transform the studied system into the proposed PONF.
- Using the theory of non-commutative rings, the delay identification problem of nonlinear time-delay systems with unknown inputs is studied in the paper [82]. Necessary and sufficient conditions are proposed to judge the identifiability of the delay, where two different cases are discussed for the dependent and independent outputs, respectively. After that, necessary and sufficient conditions are given to analyze the causal and non-causal observability for nonlinear time-delay systems with unknown inputs.

- In the paper [58], we investigate the stabilization of a linear plant subject to network constraints, partial state knowledge and time varying bounded parameter uncertainties. An event-triggered version of the Luenberger observer is proposed, and necessary conditions on the uncertainties are given in term of LMI's to enable output-based stabilization under different triggering strategies.
- The papers [47], [76] investigate an unknown input observer design for a large class of linear systems with unknown inputs and commensurate delays. A Luenberger-like observer is proposed by involving only the past and actual values of the system output. The required conditions for the proposed observer are considerably relaxed in the sense that they coincide with the necessary and sufficient conditions for the unknown input observer design of linear systems without delays.
- The paper [71] deals with the problem to estimate some states of a multi-output nonlinear dynamical system which is partially observable. To address this problem, this paper provides a set of geometrical conditions that guarantee the existence of a change of coordinates which decomposes the studied nonlinear dynamical system into two dynamical subsystems, where the first one is of the well-known output injection form. This transformed form allows us to design a simple reduced-order (Luenberger-like) observer to estimate the observable state.

6.5. Model-Free Control

- In the paper [41] the Universal Integral Control, introduced by H.K. Khalil, is revisited by employing mollifiers instead of a high-gain observer for the differentiation of the output signal. The closed loop system is a classical functional differential equation with distributed delays on which standard Lyapunov arguments are applied to study the stability. Low-pass filtering capability of mollifiers is demonstrated for a high amplitude and rapidly oscillating noise.
- The paper [64] proposes an universal adaptive control structure for robot manipulators, without knowing the dynamic model of the system, as well it is robust to corrupt payload change and initial conditions.
- In [66], the control design of an artificial pancreas, a hot research topic in diabetology, is tackled via the newly introduced model-free control and its corresponding "intelligent" proportional controller, which were already quite successful in many concrete and diverse situations. It results in an insulin injection for type 1 diabetes which displays via constant references a good nocturnal/fasting response, but unfortunately a poor postprandial behavior due to long hyperglycemia. When a variable reference is introduced, that switches between a constant one, when glycemia is more or less normal or moderate, and an exponential decay reference path, when a high glycemia rate indicates a meal intake, the results in silico, which employ real clinical data, become excellent. We obtain a bolus-shaped insulin injection rate during postprandial phases. The hyperglycemic peaks are therefore lowered a lot.

6.6. Sliding Mode Control And Estimation

- The paper [22] addresses the problem of oscillatory failure case detection in the electrical flight control system of a generic commercial airplane. A non-homogeneous differentiator is first used to provide accurate derivatives in noisy environment and fast convergence time. In this study case, fault detection is addressed in the unknown input estimation issue for fault reconstruction with the same evaluation techniques currently employed in Airbus A380 airplanes. Performance and robustness of the developed monitoring strategy are assessed using a high-fidelity Airbus benchmark and a parametric test campaign for the flight scenarios defined in the EU-FP7 ADDSAFE project.
- A new sliding mode control approach is introduced in [26] with the dedicated mathematical tools. A time-delay modification/approximation of sign function is proposed in [57], and it is shown that by substituting this new "sign" realization in the conventional sliding mode algorithms the main advantages of the sliding mode tools are preserved (like rejection of matched disturbances and hyper-exponential convergence), while the chattering is reduced.

- The article [34] proposes a convex optimization approach for the design of relay feedback controllers. The case of linear systems is studied in the presence of matched perturbations. The system input is a generalized relay that may take values in a finite set of constant vectors. A simple design method is proposed using Linear Matrix Inequalities (LMIs).
- In the note [53] we study the effect of an implicit Euler time-discretization method on the stability of the discretization of a globally fixed-time stable, scalar differential inclusion representing a simple nonlinear system with a set-valued signum controller. The controller nonlinearity is a cubic term and it is shown that the fully-implicit method preserves the global Lyapunov stability property of the continuous-time system, contrarily the explicit discretization which does not. It allows to obtain finite-time convergence to the origin when the plant is undisturbed, while the cubic term provides the hyper-exponential convergence rate.
- The problem of finite-time stabilization of multi-input linear system by means of sliding mode relay feedback is considered in [68]. A new control design procedure, which combines convex embedding technique with implicit Lyapunov function method, is developed. The issues of practical implementation of the obtained implicit relay feedback are discussed. Theoretical result is supported by numerical simulation.

6.7. Non-Linear, Sampled-Data And Time-Delay Systems

- The method of Implicit Lyapunov-Krasovski Functional (ILKF) for stability analysis of time-delay systems is introduced in [39]. Theorems on Lyapunov, asymptotic, (hyper) exponential, finite-time and fixed-time stability analysis using ILKF are presented. The hyper exponential stabilization algorithm for a time-delay system is presented.
- A recent generalization of the classical ISS theory to multistable systems is presented in [17]. Based on it a robust synchronization protocols with respect to a compact invariant set of the unperturbed system are designed in [14], [49].
- The paper [21] deals with the design of an active fault-tolerant control strategy based on the supervisory control approach technique for linear time invariant MIMO systems affected by disturbances, measurement noise, and faults. From a bank of Luenberger observers that plays the role of a fault detection and isolation scheme, the supervisory algorithm selects the suitable fault-tolerant controller by means of a hysteresis-based switching mechanism based on the method proposed in this paper.
- In [30] motivated by the problem of phase-locking in droop-controlled inverter-based microgrids with delays, the recently developed theory of input-to-state stability (ISS) for multistable systems is extended to the case of multistable systems with delayed dynamics. Sufficient conditions for ISS of delayed systems are presented using Lyapunov-Razumikhin functions. It is shown that ISS multistable systems are robust with respect to delays in a feedback [55].
- The work [31] aims at enlarging the sampling intervals in several state feedback control situations by designing a sampling map in the state space. For linear time invariant (LTI) systems with state-bounded perturbations this guarantees exponential stability with a chosen decay-rate. The approach is based on linear matrix inequalities (LMIs) obtained thanks to Lyapunov-Razumikhin stability conditions and convexification arguments. Then, the obtained results are extended to design the sampling map in three dynamic sampling control situations: event-triggered control, self-triggered control, and state-dependent sampling.
- In the paper [43] the problem of discrete and continuous state estimation for a class of uncertain switched LPV systems is addressed. Parameter identification techniques are applied to realize an approximate identification of the scheduled parameters of a switched LPV system with certain uncertainties and/or disturbances. A discrete state estimation is achieved using the parameter identification. A Luenberger-like hybrid observer, based on discrete state information and LMIs approach, is used for the continuous state estimation.

- The paper [70] contributes to the exponential stability analysis for impulsive dynamical systems based on a vector Lyapunov function and its divergence operator. The method relies on a 2D time domain representation. The results are applied to analyze the exponential stability of linear impulsive systems based on LMIs.

6.8. Networked Systems

- The problem of phase synchronization for a population of genetic oscillators (circadian clocks, synthetic oscillators, etc.) is considered in the paper [13]. The proposed analysis approach is based on the Phase Response Curve model of an oscillator. The performance of the obtained solutions is demonstrated via computer experiments for two different models of circadian/genetic oscillators.
- The paper [23] focuses on the design of fixed-time consensus for first order multi-agent systems with unknown inherent nonlinear dynamics. A distributed control protocol, based on local information, is proposed to ensure the convergence of the tracking errors in finite time. Some conditions are derived to select the controller gains in order to obtain a prescribed convergence time regardless of the initial conditions.
- The problem of phase regulation for a population of oscillating systems is considered in [24] based on a Phase Response Curve (PRC) model of an oscillator. The problem of phase resetting for a network of oscillators is solved by applying a common control input. Performance of the obtained solutions is demonstrated via computer simulation for three different models of circadian/neural oscillators.

6.9. Applications

- The problem of avoiding obstacles while navigating within an environment for a Unicycle-like Wheeled Mobile Robot (WMR) is of prime importance in robotics. The work [32] solves such a problem proposing a perturbed version of the standard kinematic model able to compensate for the neglected dynamics of the robot. The effectiveness of the solution is proved, supported by experiments and finally compared with the Dynamic Window Approach (DWA) to show how the proposed method can perform better than standard methods. The paper [60] presents a decentralized solution to control a leader-follower formation of unicycle wheeled mobile robots allowing collision and obstacle avoidance. The work [62] solves the obstacle avoidance problem extending the Potential Field (PF) method for a mobile robot. The usual definition of the PF has been modified to have a field which is continuous everywhere. It is shown that the system has an attracting equilibrium at the target point, repelling equilibriums in the centers of the obstacles and saddle points on the borders. Those unstable equilibriums are avoided capitalizing on the established Input-to-State Stability (ISS) property of this multi-stable system. To escape a local minima this work makes the most of ISS property that is not lost for perturbations. And for small properly designed disturbances the global attractivity of the target point is proved.
- The paper [63] investigates the behavior of central Jacobi differentiator in robot identification applications. It is applied to compute acceleration from noisy position measurements. Its frequency domain property is analyzed via a finite impulse response (FIR) filter point of view, indicating clearly the differentiators performance. Two revolute joints planar robot parameter identification is done. Comparisons between the Jacobi differentiator and the Euler differentiation combined with Butterworth filter are drawn.
- In [50] the velocity of valve movement activity is estimated using three different differentiation schemes: an algebraic-based differentiator method, a non-homogeneous higher order sliding mode differentiator and a homogeneous finite-time differentiator. We demonstrate that this estimated velocity can be used for water quality monitoring as the differentiators can detect very rapid change in valve movements of the oyster population resulting from some external stimulus or common input.

- In the paper [15] the measurements of valve activity in a population of bivalves under natural environmental conditions (16 oysters in the Bay of Arcachon, France) are used for a physiological model identification. A nonlinear auto-regressive exogenous (NARX) model is designed and tested. Through this study, it is demonstrated that the developed dynamical model of the oyster valve movement can be used for estimating normal physiological rhythms of permanently immersed oysters and can be considered for detecting perturbations of these rhythms due to changes in the water quality, i.e. for ecological monitoring.
- Spawning observations are important in aquaculture and biological studies, and until now, such a detection is done through visual analysis by an expert. Using measurements of valve activity (i.e. the distance between the two valves) in populations of bivalves under natural environmental condition (16 oysters in the Bay of Arcachon, France, in 2007, 2013 and 2014), algorithms for an automatic detection of the spawning period of oysters are proposed in the paper [16], [51]. The fault detection method presented in the paper can also be used to detect complex oscillatory behavior which is of interest to control engineering community.
- The work presented in the paper [33] is undertaken within the European FP7 funded Advanced Fault Diagnosis for Sustainable Flight Guidance and Control (ADDSAFE) project. It proposes new fault detection and fault diagnosis techniques that could significantly help developing environmentally-friendlier aircraft. LPV model-based fault detection schemes are proposed and compared for robust and early detection of faults in aircraft control surfaces servo-loop. The proposed methodologies are based on a slight modification of the H_∞/H_- LPV optimization techniques for systems modelled in, first polytopic manner, second linear fractional representation fashion. It is shown that the proposed fault detection schemes can be embedded within the structure of in-service monitoring systems as a part of the Flight Control Computer software. Several important examples on model and signal based fault detection in aircraft Electrical Flight Control System are studied in [80].
- For analyzing the transients of induction heating systems, time-dependent phasor transformations were proposed so far in the literature. Applying these transformations to a linear R, L, C circuit equations leads to differential equations in the complex domain from which equivalent circuits modeling the envelopes of sinusoidal waveforms were derived. The work [46] proposes a phasor transformation which is based on fictitiously replacing the real voltage and current signals of a system by complex ones. It leads to transformed system equations in the real domain where instantaneous amplitudes, phases and frequencies appear explicitly, which makes the transformed equations suitable for the feedback control design. The methodology is applied to a parallel induction heating system in order to design a sliding mode controller.
- The problem of air-to-fuel ratio regulation for a direct injection engine is addressed in [54]. A LPV model of the engine is used, for which an interval observer is designed. The interval observer is applied for the model validation and control synthesis. The results of design are confirmed by implementation.
- Modular Robot Manipulators are user-configurable manipulators which provide rapid design and inexpensive implementation. To be easy-use, smart actuators embedded with position input and position feedback controller are adopted, these local controllers render the manipulators position controlled, but also result in limited performance and precision. The paper [72] targets the case that the built-in controller does not provide desirable precision for set-point regulation. Firstly a joint-level model is established, of which the nominal model can be identified with derivative observer based on the position feedback, then an auxiliary adaptive controller coping with parametric uncertainty is proposed which leads to an error close to zero, a switching control strategy is introduced considering the actuator saturation. The paper [73] addresses the set-point control of actuators integrated with built-in controller, which presents steady-state error (SSE) under certain load. To eliminate the SSE, a model of the actuator-plus-controller system is established and identified, a switched adaptive controller is developed to work with the embedded one, considering the physical constraints, a switching control strategy is proposed. The proposed algorithms are implemented on a 5-DOF modular manipulator, with comparison to classic integral controller.

- The communication [74] is devoted to a comparison between various meteorological forecasts, for the purpose of energy management, via different time series techniques. The first group of methods necessitates a large number of historical data. The second one does not and is much easier to implement, although its performances are today only slightly inferior. Theoretical justifications are related to methods stemming from a new approach to time series, artificial neural networks, computational intelligence and machine learning.
- ALINEA is a well known ramp metering closed-loop control the aim of which is to improve highway traffic. The report [84] shows that ALINEA may be slightly modified in order to be efficiently implemented without any need of crucial time-varying quantities, like the critical density and the free-flow speed, which are most difficult to estimate correctly online.
- For malaria patients, a usual observation problem consists in estimation of sequestered parasites *Plasmodium falciparum* from measurements of circulating ones. The model of an infected patient is rather uncertain, and for all rates (death, transition, recruitment and infection) in the model it is assumed that only intervals of admissible values are given. In addition, the measurements of the concentration of circulating parasites are subjected by a bounded noise, while some parameters, like the rate of infection of blood cells by merozoites, are completely unknown and highly time-varying. In order to evaluate the concentration of sequestered parasites, an interval observer is designed in [85], which provides intervals of admissible value for that concentration, with the interval width proportional to the model uncertainty.

POEMS Project-Team

6. New Results

6.1. Wave propagation in non classical media

6.1.1. Modal analysis of electromagnetic dispersive media

Participants: Anne-Sophie Bonnet-Ben Dhia, Christophe Hazard.

Except in vacuum, the velocity of electromagnetic waves generally depends on the frequency. This dispersion plays in particular a vital role in situations where the effective index takes values below unity or negative, which happens with metamaterials or plasmonic devices. However, most of the studies in this domain are considering only the time-harmonic regime, forgetting dispersion, which leads to apparent paradoxes. We have elaborated a project, in collaboration with the Institut Fresnel in Marseille. Our objective is to gather physical and mathematical points of view to explore a frequency-to-time approach for dispersive media. This approach is based on a general technique which allows to hide dispersion in an augmented formulation of Maxwell's equations. Using this tool, our aim is first to carry the spectral analysis of dispersive systems, take advantage of this analysis to predict the time-dependent behaviour of dispersive systems, then design adapted numerical methods for their simulation and finally confirm predictions by real experiments. To begin with, during the internship of Bilal Yezza, a toy problem has been studied, where the presence of accumulation points in the spectrum is due to the dispersion. This project has been submitted to the ANR for the second year and has already led to preliminary common works and discussions, in particular during the workshop *Leaky days* organized by Christophe Hazard in Palaiseau in June 2015.

6.1.2. Perfectly Matched Layers in plasmas and metamaterials

Participants: Eliane Bécache, Patrick Joly, Maryna Kachanovska, Valentin Vinoles.

We work on the stability of Generalized Perfectly Matched Layers (GPMLs) in dispersive media for which classical PMLs are in general unstable. These new PMLs involve, in addition to the absorption parameter $\sigma \geq 0$, a real valued rational function of the frequency $\psi(\omega)$. We first worked on isotropic media and derived, using Fourier analysis methods, a necessary and sufficient condition on the function $\psi(\omega)$ for the stability of the PML model. This result has been presented in several conferences and used to design new stable PMLs for negative index metamaterials and uniaxial anisotropic plasmas (even though this last model is anisotropic, the anisotropy has a structure that permits a special decomposition of vector fields that give a new equivalent model adapted for our GPMLs).

We are currently working on the generalization of this analysis to a class of anisotropic dispersive models using a different approach based on Laplace transform in time.

However, this theory does not apply to more general cold plasma models that we wish to treat. Finding good PMLs in this case still remains a challenging open question. Several attempts, such as radial PMLs (which we discussed about with our visitor Martin Halla from TU Wien), have failed.

6.2. Wave propagation in heterogeneous media

6.2.1. Homogenization of layered media

Participant: Jean-François Mercier.

Metamaterials have revived interest in the theory of homogenization techniques because some standard techniques, based on the Ross Nicholson-Weir method, can lead to unphysical effective parameters, since depending on the incident wave. In collaboration with Agnès Maurel and Abdelkader Ourir from the Langevin Institut and Simon Felix from the LAUM, we have proposed more suitable homogenization methods to describe wave propagation in artificial environments, by considering homogenization of sliced media. When the medium is structured at a sub-wavelength scale, it can be described as a simpler equivalent medium, homogeneous and anisotropic, with a tensor mass density and an effective modulus of elasticity. We considered two cases:

- for a propagating incident wave, we obtained the diffusion properties of the medium and we have shown that the effective medium correctly captures the acoustic properties of the real medium.

- however, in the real problem, evanescent waves are generated and if one of them is resonant, the properties of transmission and reflection of the incident wave are changed: this happens for the electromagnetic waves (Wood anomalies, "spoof plasmon"). To capture these resonance effects, we have considered evanescent incident waves. We then showed that the homogenization predicts the dispersion curves of the resonant waves: in the homogenized problem, they correspond to guided waves by the anisotropic layer.

6.2.2. *High order transmission conditions between homogeneous and homogenized periodic half-spaces*

Participants: Sonia Fliss, Valentin Vinales.

This work is a part of the PhD of Valentin Vinales, and is done in collaboration with Xavier Claeys (LJLL, Paris VI). It is motivated by the fact that classical homogenization theory poorly takes into account interfaces, which is particularly unfortunate when considering negative materials, because important phenomena arise precisely at their surface (plasmonic waves for instance). To overcome this limitation, we want to construct high order transmission conditions. Using matched asymptotics, we have treated the case of a plane interface between a homogeneous and a homogenized periodic half space. The analysis is based on an original combination of Floquet-Bloch transform and a periodic version of Kondratiev techniques. The obtained conditions involve Laplace- Beltrami operators at the interface and require to solve *cell problems* in infinite strips. The numerical computations are based on specific transparent conditions for periodic media. The error analysis and the numerical study are on-going works.

6.2.3. *Scattering by small heterogeneities*

Participants: Patrick Joly, Simon Marmorat.

Simon Marmorat has defended his thesis, done in collaboration with the CEA-LIST and with Xavier Claeys (LJLL, Paris VI). The goal was to develop an efficient numerical approach to simulate the propagation of waves in concrete, which is modelled as a smooth background medium, with many small embedded heterogeneities. To do so, one has proposed two reduced models relying on the asymptotic analysis of the problem with respect to the (small) size of the heterogeneities. The first model looks like a fictitious domain method in which the analysis of the near field (closed to the heterogeneities) is exploited. The second one is a method of auxiliary sources, based on the analysis of the far field (far from the heterogeneities). Rigorous error estimates have been established. From the numerical point of view, some points, related to the Galerkin enrichment of standard finite element methods, still need to be completed.

6.2.4. *Effective boundary conditions for strongly heterogeneous thin layers*

Participants: Mathieu Chamaillard, Patrick Joly.

This topic is the object of the PhD of Mathieu Chamaillard, done in collaboration with Housseem Haddar (Inria, Defi). We are interested in the construction of effective boundary conditions for the diffraction of waves by an obstacle covered with a thin coating whose physical characteristics vary "periodically". The width of the coating and the period are both proportional to the same small parameter δ .

The results obtained previously on scalar propagation models have been extended to 3D Maxwell's equations resulting in the construction of an effective condition of the form $E \times n = \delta ik \mathcal{Z}_\Gamma (n \times (H \times n))$ where the impedance operator \mathcal{Z}_Γ , a second order tangential differential operator along Γ , depends on the geometry of the obstacle and of the material properties of the coating. The analysis, which is much more involved than in the scalar case (in particular in what concerns the stability analysis), provides error estimates in $O(\delta^2)$.

The thesis will be defended in the end of January 2016.

6.3. Spectral theory and modal approaches for waveguides

6.3.1. Guided modes in ladder-like open periodic waveguides

Participants: Sonia Fliss, Patrick Joly, Khac Long Nguyen, Elizaveta Vasilevskaya.

The general objective is the study of localized modes in locally perturbed periodic media and of guided modes in periodic media with a lineic perturbation. We investigate the existence theory of such modes as well as their numerical computations.

The problem, that is investigated in the framework of the PhD thesis of E. Vasilevskaya, in collaboration with Bérandère Delourme (Paris 13 University), is the case where the propagation medium is a thin structure whose limit is a periodic graph. We exhibit situations where the introduction of a line defect into the geometry of the domain leads to the appearance of guided modes. From the theoretical point of view, the problem is studied by asymptotic analysis methods, the small parameter being the thickness of the domain, so that when the thickness of the structure is small enough, the domain approaches a graph. The spectral theory of the underlying limit operator defined in the graph plays a key role in the analysis. For 2D configurations, we have shown that for sufficiently thin structures, it suffices to reduce the width of one rung to make appear guided modes. Moreover, using matched asymptotic expansions, we have constructed asymptotic expansions at any order of the corresponding eigenvalues and guided modes. For 3D configurations, the spectral theory of the underlying limit operator was already studied. In a further step, one can expect, again by asymptotic analysis, to get corresponding existence results for the original problem, at least for sufficiently thin structures.

From a numerical point of view, the modes can be computed using non linear eigenvalue problems and specific transparent boundary conditions for periodic media. During his post-doc, Khac Long Nguyen has implemented an exact method based on Dirichlet-to-Neumann operators to compute localized modes in 2D locally perturbed periodic media or guided modes in 3D periodic media with a lineic perturbation. This was already done few years ago for waveguides configurations but here the construction of the transparent boundary conditions are much more involved.

6.3.2. Reduced graph models for networks of thin co-axial electromagnetic cables

Participants: Geoffrey Beck, Patrick Joly.

This work is the object of the PhD of Geoffrey Beck and is done in collaboration with Sébastien Imperiale (Inria, MEDISIM). The general context is the non destructive testing by reflectometry of electric networks of co-axial cables with heterogeneous cross section and lossy materials, which was the subject of the ANR project SODDA. We consider electromagnetic wave propagation in a network of thin coaxial cables (made of a dielectric material which surrounds a metallic inner-wire). The goal is to reduce 3D Maxwell's equations to a 1D like model. During the past two years, we derived and justified generalized telegraphers model for a single cable. This year, we incorporated in our model the losses due to the skin effect induced by the non perfectly conducting nature of the metallic wire. Finally using the method of matched asymptotics, we have derived and justified improved Kirchhoff conditions.

6.3.3. Multimodal methods for the propagation of acoustic and electromagnetic waves

Participant: Jean-François Mercier.

In collaboration with Agnès Maurel from the Langevin Institut and Simon Felix from the LAUM, we have developed fast multimodal methods to describe the acoustic propagation in rigid waveguides or in periodic arrays. An incident wave is scattered by penetrable inclusions or by the succession of different penetrable media separated by interfaces of any shape. The difficulties are: to take into account the modes coupling and to get modes naturally decoupled at the entrance and at the exit of the computational domain. A weak formulation of the problem provides a modal formulation taking exactly into account the matching conditions at the interfaces. A consequence is that the obtained convergence is the best convergence expected, given the regularity of the solution. After the study of isotropic cases, we have generalized this approach to the case of anisotropic media, the difficulty being to take into account a tensor in the propagation equation.

6.3.4. Plasmonic waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Camille Carvalho, Patrick Ciarlet.

This work, which is a part of the PhD of Camille Carvalho, is done in collaboration with Lucas Chesnel (Inria, Defi). A plasmonic waveguide is a cylindrical structure consisting of metal and dielectric parts. In a certain frequency range, the metal can be seen as a lossless material with a negative dielectric permittivity. The study of the modes of a plasmonic waveguide is then presented as an eigenvalue problem with a sign-change of coefficients in the main part of the operator. Depending on the values of the contrast of permittivities at the metal / dielectric interface, different situations may occur. In the "good" case, the problem is self-adjoint with compact resolvent and admits two sequences of eigenvalues tending to $+$ and $-\infty$. But when the interface presents corners, for a particular contrast range, the problem is neither self-adjoint nor with compact resolvent. In this case, Kondratiev's theory of singularities allows to build extensions of the operator, with compact resolvent. Finally, we show that the eigenvalues for one of these extensions can be computed by combining finite elements and Perfectly Matched Layers at the corners. The paradox is that a specific treatment has to be done to capture the corners singularities, even to compute regular eigenmodes.

6.4. Inverse problems

6.4.1. Quasi-Reversibility method and exterior approach for evolution problems

Participant: Laurent Bourgeois.

This work is a collaboration with Jérémie Dardé from Toulouse University. We address some linear ill-posed problems involving the heat or the wave equation, in particular the heat/wave equation with lateral Cauchy data. We have introduced several kinds of variational mixed formulations of quasi-reversibility which enable us to solve these ill-posed problems by using classical Lagrange finite elements. We have also designed a new approach called the "exterior approach" to solve inverse obstacle problems with initial condition and lateral Cauchy data for the heat/wave equation. It is based on a combination of an elementary level set method and the quasi-reversibility methods we have just mentioned. Some numerical experiments have proved the feasibility of our strategy to identify obstacles from lateral Cauchy data for the heat equation in 2D and for the wave equation in 1D. Our objective is now to focus on the wave equation in 2D. Firstly we wish to obtain a minimal value of the final time in order to ensure uniqueness of the obstacle from the lateral Cauchy data. Secondly we want to test our exterior approach numerically. We expect better results than with the heat equation.

6.4.2. Higher-order expansion of misfit functional for defect identification in elastic solids

Participants: Marc Bonnet, Rémi Cornaggia.

This work, done in the context of the PhD of Rémi Cornaggia, concerns the identification of scatterers of moderate size, modelled as elastic inhomogeneities embedded in an homogeneous elastic background medium, by time-harmonic elastodynamic measurements. Least-squares functionals, commonly used for defect identification, are expanded in powers of the small characteristic radius a of a trial inhomogeneity. This entails the expansion of the elastodynamic scattering problem, which is needed only on the support of the trial inhomogeneity and is established by means of a Lippmann-Schwinger volume integral equation. This approach generalizes, to higher orders in a , the well-known concept of topological derivative. Such expansion,

whose derivation and evaluation are facilitated by using an adjoint state, provides a basis for the quantitative estimation of flaws whereby a region of interest may be exhaustively probed at reasonable computational cost. So far, the higher-order expansion has been derived under fairly general conditions, mathematically justified, and demonstrated on simple numerical examples involving the identification of a spherical inhomogeneity in an unbounded 3D medium.

6.4.3. Complete transmission invisibility in waveguides

Participant: Anne-Sophie Bonnet-Ben Dhia.

In collaboration with Lucas Chesnel (Inria, Defi) and Sergei Nazarov (Saint-Petersburg University), we consider time harmonic acoustic problems in waveguides. We are interested in finding localized perturbations of a straight waveguide which are not detectable in the far field, as they produce neither reflection nor conversion of propagative modes. In other words, such *invisible* perturbation produces a scattered field which is exponentially decaying at infinity in the two infinite outlets of the waveguide.

In our previous contributions, we found a way to build smooth and small perturbations of the boundary which were almost invisible, in the sense that they were producing no reflexions but maybe a phase shift in transmission.

During the visit of Sergei Nazarov, we found a new approach which allows to build completely invisible perturbations in the mono-mode regime (i.e. when the frequency is chosen below the first cut-off frequency) with no phase shift in transmission. These perturbations include some kinds of thin resonators whose height is adapted to the frequency.

All our results mainly rely on asymptotic theory.

6.4.4. Energy-based cost functional for three-dimensional transient elastodynamic imaging

Participant: Marc Bonnet.

This work is a continuing collaboration with Wilkins Aquino (Duke University, USA). It is concerned with three-dimensional elastodynamic imaging by means of the modified error in constitutive relation (MECR), combining the energy norm of the constitutive residual and a more-classical L^2 norm on the measurement residuals.

We have in particular considered the case of imaging using interior data. The stationarity equations associated with the minimization of a MECR objective function, subject to the conservation of linear momentum, yields a well-posed problem coupling two elastodynamic fields, even in cases where boundary conditions are initially underspecified (making it difficult to define *a priori* a forward problem). Numerical results demonstrate the robust performance of the method in situations where the available measurement data is incomplete and corrupted by noise of varying levels.

In a separate study, elastodynamic imaging using transient data and based on time-domain solvers has been investigated. In this context, each evaluation of a time-domain MECR cost functional entails solving two elastodynamic problems (one forward, one backward), which moreover are coupled (unlike the case of L^2 misfit functionals). This coupling creates a major computational bottleneck, making MECR-based inversion difficult for spatially 2D or 3D configurations. To overcome this obstacle, we propose to (a) set the entire computational procedure in a consistent time-discrete framework that incorporates the chosen time-stepping algorithm, and (b) use an iterative successive over-relaxation-like method for the resulting stationarity equations. The resulting MECR-based inversion algorithm is formulated under quite general conditions, allowing for 3D transient elastodynamics, straightforward use of available parallel solvers, a wide array of time-stepping algorithms commonly used for transient structural dynamics, and flexible boundary conditions and measurement settings. The proposed MECR algorithm is then demonstrated on computational experiments involving 2D and 3D transient elastodynamics and up to over 500 000 unknown elastic moduli.

6.4.5. Linear Sampling Method with realistic data in waveguides

Participants: Laurent Bourgeois, Arnaud Recoquilly.

Our activities in the field of inverse scattering in waveguides with the help of sampling methods has now a quite long history. We now intend to apply these methods in the case of realistic data, that is surface data in the time domain. This is the subject of the PhD of Arnaud Recoquilly. It is motivated by Non Destructive Testing activities for tubular structures and is the object of a partnership with CEA List (Vahan Baronian).

Our strategy consists in transforming the time domain problem into a multi-frequency problem by the Fourier transform. This allows us to take full advantage of the established efficiency of modal frequency-domain sampling methods. We have already proved the feasibility of our approach in the 2D acoustic case. In particular, we have shown how to optimize the number of sources/receivers and the distance between them in order to obtain the best possible identification result. The next steps consist in extending such an approach to the elastic case and trying it experimentally, that is with real data. Experiments will be carried in CEA.

6.5. Integral equations

6.5.1. Fast BEM solvers based on \mathcal{H} -matrices for 3-D frequency-domain elastodynamics

Participants: Stéphanie Chaillat, Patrick Ciarlet, Luca Desiderio.

The main advantage of the Boundary Element Method (BEM) is that only the domain boundaries are discretized leading to a drastic reduction of the total number of degrees of freedom. In traditional BE implementation the dimensional advantage with respect to domain discretization methods is offset by the fully-populated nature of the BEM coefficient matrix. Using the \mathcal{H} -matrix arithmetic and low-rank approximations (performed with Adaptive Cross Approximation) it is possible to derive fast iterative and direct solvers for the BEM system. We extend the method to 3-D frequency-domain elastodynamics. To this end, the Adaptive Cross Approximation is adapted to deal with vectorial problems. To validate the accuracy of the solution of the LU based direct solver, we derive an error estimate. Finally, we check numerically the theoretical estimate of the storage costs. In particular, we study the efficiency of low-rank approximations when the frequency is increased. This is done in partnership with SHELL company in the framework of the PhD of Luca Desiderio.

6.5.2. OSRC preconditioner for 3D elastodynamics

Participant: Stéphanie Chaillat.

This work is done in collaboration with Marion Darbas from University of Picardie and Frédérique Le Louer from Technological University of Compiègne. The fast multipole accelerated boundary element method (FM-BEM) is a possible approach to deal with scattering problems of time-harmonic elastic waves by a three-dimensional rigid obstacle. In 3D elastodynamics, the FM-BEM has been shown to be efficient with solution times of order $O(N \log N)$ per iteration (where N is the number of BE degrees of freedom). However, the number of iterations in GMRES can significantly hinder the overall efficiency of the FM-BEM. To reduce the number of iterations, we propose a clever integral representation of the scattered field which naturally incorporates a regularizing operator. When considering Dirichlet boundary value problems, the regularizing operator is a high-frequency approximation to the Dirichlet-to-Neumann operator. For a spherical obstacle, the approximation of the DtN is a linear combination of the tangential and normal parts. The numerical efficiency of the preconditioned integral equation (i.e. the independence of the number of iterations from the mesh size and frequency) is verified for spherical obstacles, validating the concept of analytical preconditioners for 3D elastodynamics FM-BEM. For more general shapes, this approximation of the DtN is more complex to derive. As a first step, we construct and validate the approximation in the framework of the On-Surface Radiation Condition (OSRC) method.

6.5.3. A wideband Fast Multipole Method for oscillatory kernels

Participant: Stéphanie Chaillat.

This work is done in collaboration with Francis Collino. We derive a new Fast Multipole Method (FMM) based on plane wave expansions (PWFMM), combining the advantages of the low and high frequency formulations. We revisit the method of Greengard et al. (1998) devoted to the low frequency regime and based on the splitting of the Green's function into a propagative and an evanescent part. More precisely, we give an explicit formula of the filtered translation function for the propagative part, we derive a new formula for the evanescent part and we provide a new interpolation algorithm. At all steps, we check the accuracy of the method by providing error estimates. These theoretical developments are used to propose a wideband FMM based entirely on plane wave expansions. The numerical efficiency and accuracy of this broadband PWFMM are illustrated with a numerical example.

6.5.4. *Coupling integral equations and high-frequency methods for ultrasonic NDT modelling*

Participants: Marc Bonnet, Laure Pesudo.

This work, in partnership with CEA LIST and in collaboration with Francis Collino, is undertaken in the context of the PhD thesis of Laure Pesudo. Modelling ultrasonic non destructive testing (NDT) experiments simultaneously involves the scattering of waves by defects of moderate size (for which discretization-based methods such as the BEM are appropriate) and propagation over large distances (requiring high-frequency approximations). Those two types of simulation methods are therefore simultaneously needed in NDT modelling but do not lend themselves easily to coupling. The coupling approach proposed here takes advantage of the fact that the far-field asymptotic approximation of integral representation formulas (which accurately account for the scattering by defects) yields a superposition of rays (satisfying the leading-order equations arising from high-frequency asymptotics). This allows to convert incoming rays into plane waves, compute their scattering by obstacles, and convert the scattered field into rays. A defect of given shape and characteristics becomes (approximately) represented as a point-like scatterer with anisotropic reflection properties that are computed (offline) from BEM solutions of near-field problems. Using a partition of unity on the obstacle boundary allows to approximate the obstacle by a set of point-like reflectors, thereby enlarging the size of obstacles amenable to this approach. Preliminary tests on 2D scalar wave propagation problems show that sufficient far-field accuracy is achieved for wavelength-sized defects.

6.5.5. *Dynamic soil-structure interaction*

Participants: Marc Bonnet, Stéphanie Chaillat, Zouhair Adnani.

This work, undertaken in the context of the PhD thesis of Zouhair Adnani (CIFRE partnership with EDF), concerns the simulation of dynamic soil-structure interaction (SSI) in connection with seismic assessment of civil engineering structures. The main goal is to formulate, implement, and evaluate on realistic test examples, a computational strategy that combines the fast multipole integral equation method for elastic wave propagation in unbounded regions (COFFEE FMM-accelerated BEM solver), and finite elements for modelling civil engineering structures and neighboring soil regions (the EDF in-house code Code_Aster). In a preliminary phase, the evaluation of transient elastodynamic responses via the Fourier synthesis of frequency-domain solutions computed using COFFEE (see Section 5.1) has been studied on several test problems, achieving substantial improvements of computational efficiency for this component of SSI analysis.

6.5.6. *Volume Integral Formulations*

Participant: Marc Bonnet.

Volume integral equations (VIEs), also known as Lippmann-Schwinger integral equations, arise naturally when considering the scattering of waves by penetrable, and possibly heterogeneous, inhomogeneities embedded in a homogeneous background medium. In contrast with the vast existing literature on boundary integral equations, comparatively few studies are available regarding the mathematical properties of VIEs. In this work, we investigate the solvability of VIE formulations arising in elastodynamic scattering by penetrable obstacles. The elasticity tensor and mass density are allowed to be smoothly heterogeneous inside the obstacle and may be discontinuous across the background-obstacle interface, the background elastic material being homogeneous. Both materials may be anisotropic, within certain limitations for the background medium. The VIE associated with this problem is derived, relying on known properties of the background fundamental tensor.

To avoid difficulties associated with existing radiation conditions for anisotropic elastic media, we propose an alternative definition of the radiating character of transmission solutions. The unique solvability of the volume integral equation (and of the scattering problem) is established. For the important special case of isotropic background properties, our definition of a radiating solution becomes equivalent to the classical Sommerfeld-Kupradze radiation conditions.

6.6. Domain decomposition methods

6.6.1. *Transparent boundary conditions with overlap in unbounded anisotropic media*

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Antoine Tonnoir.

We are interested in acoustic or elastic wave propagation in time harmonic regime in a 2D or 3D medium which is a local perturbation of an infinite anisotropic homogeneous medium. We investigate the question of deriving a formulation which is suitable for numerical computations. This question is difficult due to the anisotropy of the surrounding medium. Our approach consists in coupling several plane-waves representations of the solution in half-spaces surrounding the defect with a FE computation of the solution around the defect. The difficulty is to ensure that all these representations match, in particular in the infinite intersections of the half-spaces. It leads to a formulation which couples, via integral operators, the solution in a bounded domain including the defect and its traces on the edge of the half-planes. We have shown that this formulation has good properties from theoretical and numerical points of view.

6.6.2. *Electromagnetic scattering by objects with multi-layered dielectric coatings*

Participants: Patrick Joly, Matthieu Lecouvez.

The PhD thesis of Matthieu Lecouvez, undertaken in collaboration with the CEA-CESTA and Francis Collino, has been defended in July. It concerned the diffraction of time harmonic electromagnetic waves by perfectly conducting objects covered by multi-layered (possibly thin) dielectric coatings. This idea was to use a domain decomposition method in which each layer would constitute a subdomain. The transmission conditions between the subdomains involve some specific impedance operators in order to achieve a geometric convergence of the method (compared to the slow algebraic convergence obtained with standard Robin conditions). This year, the theoretical aspects of our work have been completed and are the object of an article in preparation.

6.6.3. *Domain Decomposition Methods for the neutron diffusion equation*

Participants: Patrick Ciarlet, Léandre Giret.

Studying numerically the steady state of a nuclear core reactor is expensive, in terms of memory storage and computational time. In its simplest form, one must solve a neutron diffusion equation with low-regularity solutions, discretized by finite element techniques, totaling millions of unknowns or more, within a loop. Iterating in this loop allows to compute the smallest eigenvalue of the system, which determines the critical, or non-critical, state of the 3D core configuration. This problem fits within the framework of high performance computing so, in order both to optimize the memory storage and to reduce the computational time, one can use a domain decomposition method, which is then implemented on a parallel computer. The definition of an efficient DD method has been addressed for conforming meshes prior to the PhD research of Léandre Giret. The development of non-conforming, hence more flexible, DD methods has recently been finalized. The optimization of the eigenvalue loop is under way. The current research also focuses on the numerical analysis of the full suite of algorithms to prove convergence for the complete multigroup SPN model (which involves coupled diffusion equations). Realistic computations will be carried out with the APOLLO3 neutronics code.

This topic is developed in partnership with CEA-DEN (Erell Jamelot).

6.7. Aeroacoustics

6.7.1. Time-harmonic acoustic scattering in a vortical flow

Participants: Antoine Bensalah, Patrick Joly, Jean-François Mercier.

This activity is done in the framework of the PhD of Antoine Bensalah, in partnership with Airbus Group. We study the time-harmonic acoustic radiation in a fluid in a general flow which is not curl free, but has restricted vortical areas. The objective is to take into account the complicated coupling between acoustics and hydrodynamics. The Galbrun approach developed previously in 2D is too expensive in terms of degrees of freedom for 3D simulations. As an alternative, we propose to consider instead the Goldstein equations, which are vectorial only in the vortical areas and remain scalar elsewhere.

We have proved that the Goldstein equations are well-posed in a domain Ω for a potential flow, or for a vortical flow if the flow is Ω -filling (each point of Ω is reached by a streamline coming from the inflow boundary in a finite time). A non Ω -filling flow corresponds to the presence of recirculations areas and we have shown that, for such flows, some of the closed streamlines can be resonant. To study deeper this phenomenon, we focused on the case of a rotating flow in an annular geometry. We proved that outside the set of resonance frequencies, the radiation problem is well-posed. Work is under progress to determine the solution on a resonant streamline.

6.7.2. Propagation of solitons through Helmholtz resonators

Participant: Jean-François Mercier.

With Bruno Lombard (Laboratoire de Mécanique et Acoustique of Marseille), we studied the propagation of an acoustic solitary wave in a 1D waveguide connected to a periodic array of Helmholtz resonators. Starting from a model of the literature, obtained by approximations, our goal was to provide a numerical modeling, which validates (or not) the underlying model and the assumptions. The model consists of two coupled equations evolution: a nonlinear PDE describing acoustic waves (similar to the Burgers equation), and a linear ODE describing oscillations in the Helmholtz resonators. We have developed a numerical method based on two main ingredients: a diffusive representation of fractional derivatives and a splitting method applied to the evolution equations. The numerical scheme has been validated by comparison with exact solutions. The properties of non-linear solutions have been investigated numerically.

In collaboration with O. Richoux of the LAUM, this work has been extended, comparing to experimental results. Adjustments had to be made, the attenuation of the numerical model being weaker than that observed experimentally. To remedy this, we have incorporated some attenuation mechanisms that we had neglected. One consequence of these additions is that a more sophisticated numerical method had to be developed. A good agreement has been found with experimental results.

QUANTIC Project-Team

6. New Results

6.1. Entanglement between stationary and propagating modes

Participants: B. Huard and F. Mallet.

The results of this section were published in [14].

Entanglement being instrumental in quantum machines, we have shown how a Josephson mixer can generate and distribute entangled microwave radiations on separated transmission lines and different frequencies by spontaneous parametric down-conversion in 2012. Using two Josephson mixers, we have provided the first demonstration of entanglement between spatially separated propagating fields in the microwave domain. Therefore, a new variety of entangled states, the so-called EPR states (after Einstein, Podolsky and Rosen), which are encoded on continuous variables, is now available in this frequency range.

In 2015, we have shown that it could constitute the central component of a potential quantum network based on continuous-variable entanglement. The device essentially acts as a regular mixer performing frequency conversion but without adding extra noise. Used as a switch, it is able to open and close the coupling to a high-quality factor cavity in a time-controlled way. We have demonstrated how this feature leads to a new kind of quantum memory. Coupled to its ability to generate entanglement, we have demonstrated the time-controlled generation, storage and on-demand release of an entangled state, which is the prerequisite for the node of a quantum network.

Several implementations of quantum memories for microwave radiation have been realized in the past few years. In order to store the state of microwave signals, some use spin ensembles [81], [130], [71], or mechanical oscillators [98], while others use superconducting cavities with tunable input coupling [124], [102]. Our own implementation is sketched in Fig. 3 b, where the Josephson Mixer allows an on-demand access to the long lived 3D cavity based on noiseless frequency conversion. Its main advantage consists in the ability to generate entanglement between the memory and the output port.

Noiseless frequency conversion is another regime of the Josephson mixer. The frequency of the pump tone is now chosen to be at the difference between the frequencies of the modes \hat{a} and \hat{b} , $\Omega = |\omega_a - \omega_b|$. In the rotating frame, the effective Hamiltonian reduces to a beam-splitter Hamiltonian with an implicit frequency conversion:

$$H = \hbar\chi(\hat{a}^\dagger\hat{b}\hat{c} + \hat{a}\hat{b}^\dagger\hat{c}^\dagger).$$

The elementary process corresponds to the conversion of photons between the mode a and b mediated by the pump at a rate $\chi|\langle\hat{c}\rangle|$ as sketched in Fig. 3 c. Therefore, the noiseless frequency conversion generates a coupling between the long lived cavity mode \hat{b} and the propagating modes at the input of mode \hat{a} . This pump field can then be varied in time to switch on and off the coupling.

A first measurement consists in the capture, storage and retrieval of a microwave pulse. The protocol is quite simple, we turn the pump tone on when the incoming pulse reaches the memory input. The signal pulse has been designed such that it is optimally absorbed by the memory. The pump tone is turned off after the absorption and turned back on at a later time τ to retrieve the pulse in the transmission line. The measured output amplitude in time shown in Fig. 3 d demonstrate that this protocol can be performed with a great efficiency for a few microseconds.

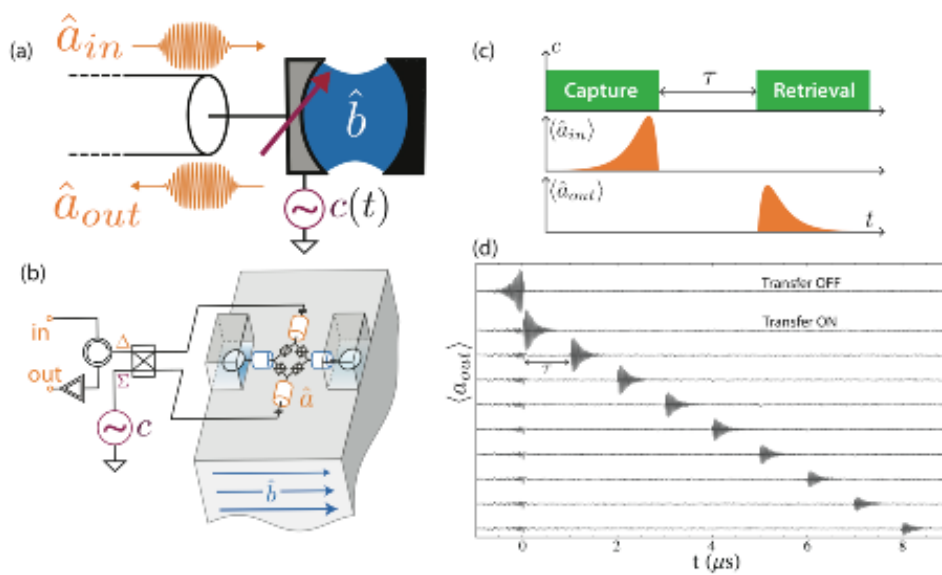


Figure 3. (a) Simplified schematics of the quantum memory. When the pump is driven at $\Omega = |\omega_a - \omega_b|$, the JRM behaves as a beam splitter with an implicit frequency conversion whose transparency depends on the pump amplitude. (b) Schematics of the device. The core of the device is similar to the usual design [107] excepted that one of the two transmission lines is replaced by a superconducting 3D cavity that defines the memory mode. (c) Protocol of the capture, storage and release of an incoming microwave pulse. (d) Measured output amplitude as a function of time. In the first trace, the pump is always turned off and the measured amplitude corresponds to the reflected incoming pulse. In the following traces, the pump is turned on and varied in time as indicated in (c). The storage time is varied from 0 μs to 8 μs .

However, the unique ability of this device lies in the possibility to combine this storage operation with the entanglement generation demonstrated previously. A second measurement consists in the generation, storage and characterization of an EPR state distributed between the memory and the transmission line. The protocol is sketched in Fig. 4 b. The pump is first applied at $\Omega = \omega_a + \omega_b$ to generate an EPR state shared between the memory and the propagating mode. The propagating mode complex amplitude is measured and at a later time, the pump is turned on again at $\Omega = |\omega_a - \omega_b|$ to activate the noiseless conversion. The memory mode is then retrieved in the transmission line and its complex amplitude is measured. By analyzing the cross-correlations between these two measurements, we have been able to show that the memory preserves the entanglement of the EPR state. Furthermore, the contours of the EPR state Wigner function have been inferred from this correlation measurement (Fig. 4 c) and the entanglement quantified.

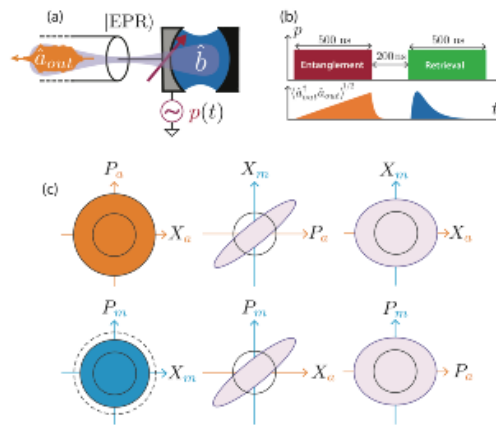


Figure 4. (a) When the pump is shined at $\Omega = \omega_a + \omega_b$, an EPR state is distributed between the transmission line and the memory. (b) Protocol for the entanglement distribution, storage and retrieval. (c) Contour of the marginal Wigner distributions reconstructed from the correlation measurements corresponding to the protocol (b).

6.2. Wideband Josephson mixer

Participants: B. Huard and F. Mallet.

The results of this section were published in [22].

For nearly a decade, the superconducting circuits community develops microwave amplifiers in the quantum regime, i.e. adding only a noise comparable to the vacuum fluctuations of the signal. We participated in this effort in 2012 [107] by adding frequency tunability to the only non-degenerate existing amplifier: the Josephson Parametric Converter (JPC) invented by the group of Michel Devoret at Yale.

However, this amplifier showed the defect of being limited to a few MHz bandwidth for a gain of 20 dB and a dynamic range (maximum input power before changing the gain) capable of amplifying signals typical of circuit-QED. We conducted a theoretical study to understand the various constraints involved in the manufacture of such an amplifier. This study has allowed us to make the first lumped element version of the JPC with bandwidth only limited by the mismatch between the characteristic impedance of the resonators and that of the transmission line.

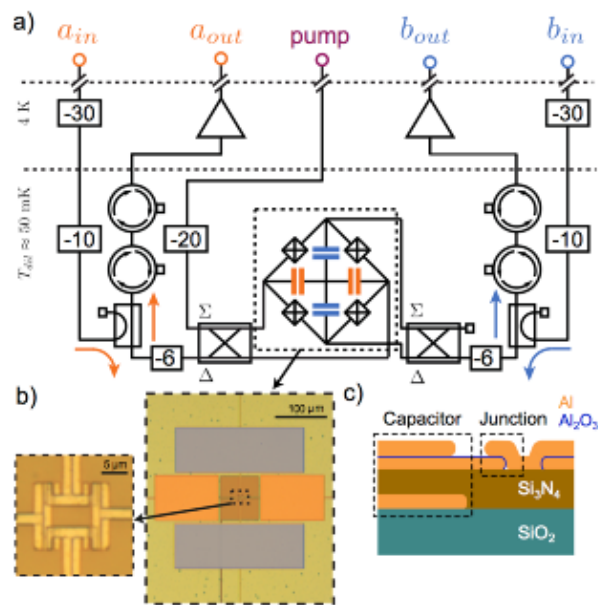


Figure 5. (a) Simplified schematic of the experimental setup. Differential a and b modes of the Josephson mixer are addressed in reflection through two 180 degree hybrid couplers. All input lines are filtered and attenuated (partially shown). Output signals are separated from input signals by a directional coupler and amplified by a low noise HEMT amplifier at 4K. (b) Optical microscope picture of the device showing the planar capacitors (right) and the Josephson junction ring (left). (c) Side view of the device. The thickness of the bottom plate of the capacitors is 35 nm and buried below 200 nm of silicon nitride, the top plate of the capacitors and the Josephson junctions are obtained by double angle deposition of 100 nm and 120 nm of aluminium with an intermediate oxidation.

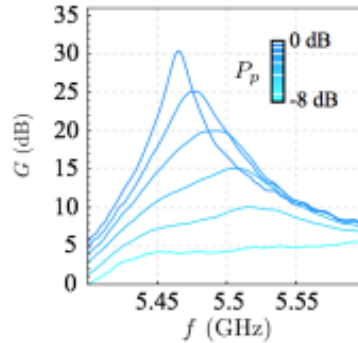


Figure 6. Gain in reflection as a function of frequency for various pump powers. The color bar encodes the pump power referred to the parametric oscillation threshold.

Finally we have measured the quantum efficiency of this amplifier and obtained almost 70%, which means that only 30% of the noise power observed at the end of line comes from technical noise while 70% is the signal, including quantum noise.

6.3. Quantum Zeno dynamics

Participants: B. Huard, L. Bretheau, P. Campagne-Ibarcq, F. Mallet.

The results of this section were published in [13].

Electromagnetic modes are instrumental for realizing quantum physics experiments and building quantum machines. Their manipulation usually involves the tailoring of their Hamiltonian in time. An alternative control scheme, called Quantum Zeno Dynamics (QZD), consists in restricting the evolution of a mode to a subset of possible states. This promising control scheme had been implemented in 2014 on atomic levels of Rb and of a Rydberg atom.

We have made the first observation of QZD of light, using superconducting circuits. By preventing the access to a single energy level, the dynamics of the field is dramatically changed. In this experiment, it was indeed possible to avoid a number of photons N , which was arbitrarily chosen between 2 and 5. Under this constraint, and starting in its ground state, a resonantly driven mode is confined to levels 0 to $N - 1$. The level occupation is then found to oscillate in time, similarly to an N -level system. Performing a direct Wigner tomography of the field reveals its non-classical features. In particular, at half period in the evolution, it resembles a "Schrödinger cat state".

In its original definition, the quantum Zeno effect corresponds to the inhibition of coherent transitions from, or to, the pointer states of a strong measurement or dissipative process. Instead of freezing the dynamics, one can restrict it to a given subspace by choosing a measurement with degenerate eigenvalues.

Similar behavior can also be induced by rapid unitary "kicks", leaving the subspace to protect unaffected. It can be understood considering a model for the original Zeno measurement as a series of coherent interactions with ancillary systems. When the interactions are strong enough, departure from the subspace is perfectly suppressed, so that the outcome of the detector is always the same. Therefore, the ancillas are all left in the same state after the interaction and they do not need to be reset. One can then enforce Zeno dynamics by performing repeatedly unitary operations controlling the state of an auxiliary degree of freedom. This amounts

to re-using the same ancilla, at the condition that the unitary evolutions are fast enough to effectively randomize the phase of coherences created with the system. In that sense, QZD is a coherent feedback, which engineers the energy level landscape of a system or its environment by coherent coupling with an ancillary degree of freedom.

In the experiment, a qubit in the resolved photon number regime plays the role of the ancillary system. A strong Rabi drive is applied on its transition conditioned on the cavity mode hosting N photons ($N = 3$ on Fig. 7). The drive hybridizes the levels $|N, g\rangle$ and $|N, e\rangle$ that repel each other. The level $|N\rangle$ is then moved out from the harmonic ladder of the cavity mode. When starting in the vacuum and applying a coherent drive at ω_r , the generated state cannot contain N photons so that it is restricted to N levels.

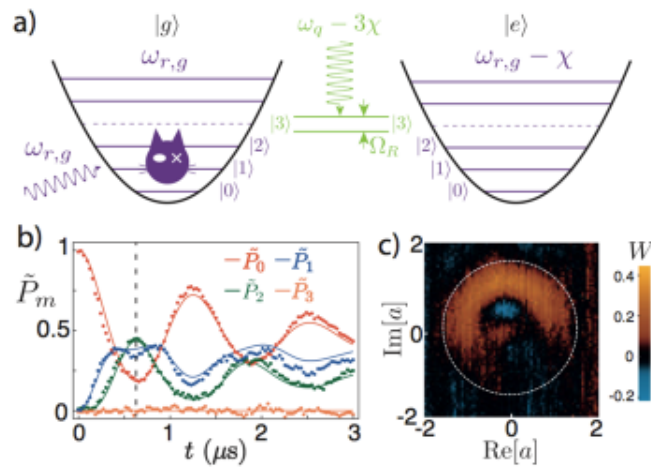


Figure 7. a) Combined energy level diagram for the qubit and cavity. By applying a strong Rabi drive on the $|3, g\rangle \longleftrightarrow |3, e\rangle$ transition, the $|2\rangle \longleftrightarrow |3\rangle$ transition of the cavity becomes off resonant at $\omega_{r,g}$. b) Oscillations of the Fock state occupation when driving the cavity mode from the vacuum and blocking $|3\rangle$. c) Wigner tomography of the field at half period of oscillation (dashed line in b). The quasi-probability density is confined within a circular barrier of radius $\sqrt{3}$ (white circle). Negativities (in blue) reveal a non classical state.

When measuring the Fock state occupation probabilities as a function of time for this effective driven N -level system, characteristic oscillations appear (see Fig. 7 b). Quantum coherence of the field is revealed by direct Wigner tomography (see Fig. 7 c). At half-period of the oscillations, fringes with negativities can be observed. This non classical state is similar to a "Schrödinger cat state", confined in phase space within a circular barrier of radius \sqrt{N} .

All these observations are well captured by a model based on N levels only. Our results demonstrate that QZD allows the direct control of the field state in its phase space. This experiment paves the way to the realization of various protocols, such as phase space tweezers, generation and protection of entanglement, and quantum logic operations.

6.4. Efficient quantum filtering for quantum feedback control

Participants: Pierre Rouchon

The results of this section were published in [23].

We discuss an efficient numerical scheme for the recursive filtering of diffusive quantum stochastic master equations. We show that the resulting quantum trajectory is robust and may be used for feedback based on inefficient measurements. The proposed numerical scheme is amenable to approximation, which can be used to further reduce the computational burden associated with calculating quantum trajectories and may allow real-time quantum filtering. We provide a two-qubit example where feedback control of entanglement may be within the scope of current experimental systems.

6.5. Adaptive low-rank approximation and denoised Monte-Carlo approach for high-dimensional Lindblad equations

Participants: Pierre Rouchon

The results of this section were published in [17].

We present a twofold contribution to the numerical simulation of Lindblad equations. First, an adaptive numerical approach to approximate Lindblad equations using low-rank dynamics is described: a deterministic low-rank approximation of the density operator is computed, and its rank is adjusted dynamically, using an on-the-fly estimator of the error committed when reducing the dimension. On the other hand, when the intrinsic dimension of the Lindblad equation is too high to allow for such a deterministic approximation, we combine classical ensemble averages of quantum Monte Carlo trajectories and a denoising technique. Specifically, a variance reduction method based upon the consideration of a low-rank dynamics as a control variable is developed. Numerical tests for quantum collapse and revivals show the efficiency of each approach, along with the complementarity of the two approaches.

This work results from a collaboration with Claude Le Bris of the Materials project-team and in the framework of the ANR-project EMAQS entitled "Evaluation and Manipulation At Quantum Scale" coordinated by Karine Beauchard from ENS-Rennes.

6.6. Stabilization of photon-number states via single-photon corrections: a first convergence analysis under an ideal set-up

Participants: Pierre Rouchon

The results of this section were published in [33].

This work presents a first mathematical convergence analysis of a Fock states feedback stabilization scheme via single-photon corrections. This measurement-based feedback has been developed and experimentally tested in 2012 by the cavity quantum electrodynamics group of Serge Haroche and Jean-Michel Raimond. Here, we consider the infinite-dimensional Markov model corresponding to the ideal set-up where detection errors and feedback delays have been disregarded. In this ideal context, we show that any goal Fock state can be stabilized by a Lyapunov-based feedback for any initial quantum state belonging to the dense subset of finite rank density operators with support in a finite photon-number sub-space. Closed-loop simulations illustrate the performance of the feedback law.

Paulo Sergio Pereira da Silva and Pierre Rouchon are participants to the Inria associate Team CDSS with principal Inria investigator, François Dufour of the Inria Team Project CQFD on the topic "Control of dynamic systems subject to stochastic jumps".

6.7. Convergence and adiabatic elimination for a driven dissipative quantum harmonic oscillator

Participants: Rémi Azouit, Alain Sarlette, Pierre Rouchon

The results of this section were published in [30].

We prove that a harmonic oscillator driven by Lindblad dynamics where the typical drive and loss channels are two-photon processes instead of single-photon ones, converges to a protected subspace spanned by two coherent states of opposite amplitude. We then characterize the slow dynamics induced by a perturbative single-photon loss on this protected subspace, by performing adiabatic elimination in the Lindbladian dynamics.

6.8. Parameter estimation from measurements along quantum trajectories

Participants: Pierre Six, Ph. Campagne-Ibarcq, Benjamin Huard, Pierre Rouchon

The results of this section were published in [34].

The dynamics of many open quantum systems are described by stochastic master equations. In the discrete-time case, we recall the structure of the derived quantum filter governing the evolution of the density operator conditioned to the measurement outcomes. We then describe the structure of the corresponding particle quantum filters for estimating constant parameter and we prove their stability. In the continuous-time (diffusive) case, we propose a new formulation of these particle quantum filters. The interest of this new formulation is first to prove stability, and also to provide an efficient algorithm preserving, for any discretization step-size, positivity of the quantum states and parameter classical probabilities. This algorithm is tested on experimental data to estimate the detection efficiency for a superconducting qubit whose fluorescence field is measured using a heterodyne detector.

6.9. Adding a single state memory optimally accelerates symmetric linear maps

Participants: Alain Sarlette

The results of this section are to be published in IEEE Trans. Automatic Control [24].

This work is exploring the context and benefits of so-called “non-Markovian” dynamics, where the dynamics implied by hidden variables modifies the behavior of an iterative procedure. Such mechanisms appear in both classical and quantum systems, and one of our future goals is to better characterize the benefits of engineered non-Markovianity in terms of stabilizing power in very constrained systems. The precise setting here is a discrete-time linear map, which is unknown except for a lower and upper bound on its eigenvalues. By adding one memory slot to each coordinate, this map can be accelerated quadratically. We prove that by adding more memory slots, this cannot be further improved. This is reminiscent of the acceleration of random walks by lifting them or by quantizing them, which we are currently exploring.

6.10. A common symmetrization framework for iterative (linear) maps

Participants: Alain Sarlette

The results of this section were presented at [29].

We review a “symmetrization” abstraction of iterative consensus algorithms, which allows to generalize them to general discrete group operations including those acting on quantum systems and on sequences of control actions. We highlight a few new applications of the framework including: consensus networks with antagonistic interactions; sub-stochastic matrix iterations; and coordinate descent on (locally) quadratic functions. The purpose is to show which types of iterative dynamics can be covered by this group-theoretic framework, and potentially operationally generalized to non-classical systems.

6.11. Deterministic hidden coordinate for a qubit under fluorescence measurement

Participants: Alain Sarlette, Pierre Rouchon

The experimentalists in the group have set up an experiment with continuous heterodyne measurement of an energy loss operator on a superconducting qubit. We have observed that in the associated mathematical model, due to the degeneracy of the diffusion operator, the resulting quantum trajectories are supported not in the entire Bloch sphere, but instead they belong to the surface of a *deterministically* evolving ellipsoid. We have entirely characterized this fact and highlighted that such behavior is not generic. A paper comparing this to the experimental data and a more general theory about deterministic evolutions in quantum stochastic differential equations are being finalized. This work has been presented at [28].

6.12. Relations between quantum walks, open quantum walks, and lifted walks: the cycle graph

Participants: Alain Sarlette

The convergence time of a random walk on a graph towards its stationary distribution is an important indication of the efficiency of random algorithms based on it. Quantum random walks have been shown to allow quadratically accelerated convergence for large graphs, at least in some cases. The famous Grover search algorithm has been shown to actually fit this framework in an abstracted setting (it is doing the opposite of a random walk: converging from the uniform distribution towards a particular identified element). Yet also with classical dynamics, simple mechanisms have been proposed which allow to quadratically accelerate the convergence with respect to a standard random walk. Some basic principles have been conjectured to cause this acceleration, basically transforming a diffusion-like behavior into a more transport-like behavior, but with remaining trail. We are working towards formally characterizing the effect of these principles, and extracting similar principles in the quantum walks. This should help identify key effects to be protected in the associated quantum algorithms. We currently have worked out the equivalence of all these accelerating settings for the simplest example of the cycle graph. Quantum coherences turn out to play no major role and a classical feedback structure can be identified. We are now working towards other graphs, where the convergence effect of quantum coherences might be hidden in propagating classical information. This work has been presented at [35].

6.13. Confining the state of light to a quantum manifold by engineered two-photon loss

Participants: Zaki Leghtas and Mazyar Mirrahimi

Physical systems usually exhibit quantum behavior, such as superpositions and entanglement, only when they are sufficiently decoupled from a lossy environment. Paradoxically, a specially engineered interaction with the environment can become a resource for the generation and protection of quantum states. This notion can be generalized to the confinement of a system into a manifold of quantum states, consisting of all coherent superpositions of multiple stable steady states. In a collaboration with the team of Michel H. Devoret at Yale university, we have confined the state of a superconducting resonator to the quantum manifold spanned by two coherent states of opposite phases and have observed a Schrödinger cat state spontaneously squeeze out of vacuum before decaying into a classical mixture. As suggested by our earlier work [93], this experiment points toward robustly encoding quantum information in multidimensional steady-state manifolds and should lead to significant hardware shortcuts for quantum error correction and fault-tolerant quantum computation.

This experimental work was published in Science [18].

6.14. Single-Photon-Resolved Cross-Kerr Interaction for Autonomous Stabilization of Photon-Number States

Participants: Zaki Leghtas and Mazyar Mirrahimi

Quantum states can be stabilized in the presence of intrinsic and environmental losses by either applying active feedback conditioned on an ancillary system or through reservoir engineering. Reservoir engineering maintains a desired quantum state through a combination of drives and designed entropy evacuation. In a collaboration with the team of Robert J. Schoelkopf at Yale university, we propose and implement a quantum reservoir engineering protocol that stabilizes Fock states in a microwave cavity. This protocol is realized with a circuit quantum electrodynamics platform where a Josephson junction provides direct, nonlinear coupling between two superconducting waveguide cavities. The nonlinear coupling results in a single photon resolved cross-Kerr effect between the two cavities enabling a photon number dependent coupling to a lossy environment. The quantum state of the microwave cavity is discussed in terms of a net polarization and is analyzed by a measurement of its steady state Wigner function.

This work was published in Physical Review Letters [15].

6.15. Characterizing entanglement of an artificial atom and a cavity cat state with Bell's inequality

Participants: Zaki Leghtas and Mazyar Mirrahimi

The Schrödinger's cat thought experiment highlights the counterintuitive concept of entanglement in macroscopically distinguishable systems. The hallmark of entanglement is the detection of strong correlations between systems, most starkly demonstrated by the violation of a Bell inequality. No violation of a Bell inequality has been observed for a system entangled with a superposition of coherent states, known as a cat state. In a collaboration with the team of Robert J. Schoelkopf at Yale university, we use the Clauser-Horne-Shimony-Holt formulation of a Bell test to characterize entanglement between an artificial atom and a cat state, or a Bell-cat. Using superconducting circuits with high-fidelity measurements and real-time feedback, we detect correlations that surpass the classical maximum of the Bell inequality. We investigate the influence of decoherence with states up to 16 photons in size and characterize the system by introducing joint Wigner tomography. Such techniques demonstrate that information stored in superpositions of coherent states can be extracted efficiently, a crucial requirement for quantum computing with resonators.

This work was published in Nature Communications [25].

RAPSODI Team

7. New Results

7.1. Design and analysis of advanced finite volumes schemes

The fact that a numerical method is able to handle nonlinear test functions in its numerical analysis is crucial in order to ensure its physical relevance, and consequently its good behavior.

In [15], C. Cancès and C. Guichard proposed a first nonlinear numerical method to solve possibly degenerate parabolic equations with anisotropy on general simplicial meshes. The nonlinear control volume finite element (CVFE) scheme is based on P1 finite elements with mass-lumping combined with a tricky upwinding of the mobilities. The method has the remarkable property of preserving the positivity of the solutions. Moreover, it ensures the decay of the physical entropy. Its convergence is proved in [15] and numerical results are exhibited. In particular, they show that the method is first order accurate in space in standard situations, but can lack robustness w.r.t. the anisotropy in some particularly unfavorable situations.

This drawback was corrected by C. Cancès and C. Guichard in [35], where a second order in space method based on the so-called VAG scheme [57] was proposed. This method is able to handle very general grids, heterogeneous data and strong anisotropy ratios. Moreover, it preserves at the discrete level the variational structure of the continuous problem, yielding the nonlinear stability of the scheme. A complete convergence analysis was performed in [35]. The numerical results presented in [35] show that the robustness default of the first nonlinear method [15] has been corrected.

In [36], C. Cancès *et al.* proposed and analyzed a nonlinear CVFE scheme for a degenerate Keller-Segel model with anisotropic and heterogeneous diffusion tensors. The scheme is based on the one proposed in [15]. The convergence of the scheme is proved under very general assumptions. Finally, some numerical experiments are carried out to prove the ability of the scheme to tackle degenerate anisotropic and heterogeneous diffusion problems over general meshes without jeopardizing the positivity of the solutions.

In [17], C. Chainais-Hillairet, A. Jüngel and S. Schuchnigg prove the time decay of fully discrete finite-volume approximations of porous-medium and fast-diffusion equations with Neumann or periodic boundary conditions in the entropy sense. The algebraic or exponential decay rates are computed explicitly. In particular, the numerical scheme dissipates all zeroth-order entropies which are dissipated by the continuous equation. The proofs are based on novel continuous and discrete generalized Beckner inequalities.

In [18], C. Chainais-Hillairet, A. Jüngel and P. Shpartko propose and analyze a numerical scheme for a spinorial matrix-diffusion model for semiconductors. The model consists of strongly coupled parabolic equations for the electron density matrix or, alternatively, of weakly coupled equations for the charge and spin-vector densities, coupled to the Poisson equation for the electric potential. The main features of the numerical scheme are the preservation of nonnegativity and L^∞ bounds of the densities and the dissipation of the discrete free energy. The existence of a bounded discrete solution and the monotonicity of the discrete free energy are proved. The fundamental ideas are reformulations using spin-up and spin-down densities and certain projections of the spin-vector density, free energy estimates, and a discrete Moser iteration. Furthermore, numerical simulations of a simple ferromagnetic-layer field-effect transistor in two space dimensions are presented.

In [32], M. Bessemoulin-Chatard and C. Chainais-Hillairet study the large-time behavior of a numerical scheme discretizing drift-diffusion systems for semiconductors. The numerical method is finite volume in space, implicit in time, and the numerical fluxes are a generalization of the classical Scharfetter-Gummel scheme which allows to consider both linear or nonlinear pressure laws. They study the convergence of approximate solutions towards an approximation of the thermal equilibrium state as time tends to infinity, and obtain a decay rate by controlling the discrete relative entropy with the entropy production. This result is proved under assumptions of existence and uniform-in-time L^∞ estimates for numerical solutions, which are then discussed.

7.2. A posteriori analysis and computational optimization

In 2015, E. Creusé *et al.* have developed a posteriori error estimators for the harmonic potential formulations of the Maxwell system, in order to simulate eddy-current problems arising in the context of quasi-static approximations. The originality of our contribution is to provide estimators with sharp bounds and explicit constants. It was achieved by solving in the same time the so-called " \mathbf{A}/φ " and " \mathbf{T}/Ω " potential formulations [38]. If this way to proceed was already known and usually used for stationary problems, the extension to harmonic ones constitutes the novelty of our contribution. It was in particular necessary to prove some superconvergence properties of additional terms. The reliability as well as the local efficiency of the derived estimator have been established without any generic constant, and numerical tests clearly illustrate their optimal behavior, from academic benchmarks to more industrial ones.

Another track to optimize the computational effort consists in refining and coarsening the model. This approach is based on the following ansatz : the more the model is complex, the more expensive are the computations. This approach was used by F. Filbet and T. Rey in [23] to simulate kinetic equations, the kinetic equations being replaced by cheaper hydrodynamic limits when it is relevant. The same idea was used in H. Mathis *et al.* [27] in order to simulate complex flows modeled by hyperbolic systems with relaxation. A rigorous error analysis of such a model adaptation procedure was performed on a simplified model by C. Cancès *et al.* in [13].

7.3. Modeling and numerical simulation of complex fluids

Recently, C. Calgaro *et al.* compared some very recent numerical schemes for the resolution of incompressible variable density flows; namely an Hybrid Finite Volume/Finite Element scheme, and a Discrete Duality Finite Volume one [34]. This work was performed in collaboration with the Inria team COFFEE (Inria Nice Sophia-Antipolis). In addition to this original and attentive comparison, our main contribution has been to improve the way to implement the Neumann boundary condition on the density, when a second-order accurate scheme is considered in space. Indeed, for some critical situations such as the simulation of Rayleigh-Taylor instabilities using unstructured meshes, this point is crucial to avoid unphysical numerical instabilities in the vicinity of the boundaries corresponding to symmetric axis. The obtained results are very promising, and constitute an important step towards the simulation of more complex models on which we are working at the moment.

In [25], M. Gisclon and I. Lacroix-Violet consider the barotropic compressible quantum Navier-Stokes equations with a linear density dependent viscosity and its limit when the scaled Planck constant vanish. Following recent works on degenerate compressible Navier-Stokes equations, we prove the global existence of weak solutions by the use of a singular pressure close to vacuum. With such singular pressure, we can use the standard definition of global weak solutions which also allows to justify the limit when the scaled Planck constant denoted by ε tends to 0.

The H-theorem, originally derived at the level of the Boltzmann nonlinear kinetic equation for a dilute gas undergoing elastic collisions, strongly constrains the velocity distribution of the gas to evolve irreversibly towards equilibrium. As such, the theorem could not be generalized to account for dissipative systems: the conservative nature of collisions is an essential ingredient in the standard derivation. The work [24] gives the first strong numerical evidences, along with a proof for a simplified model, of dissipation of the Boltzmann entropy (the so-called H-theorem) for solutions to the granular gases equation. This dissipative kinetic equation describes the non-equilibrium behavior of a gas composed of macroscopic particles, namely complex fluids such as avalanches, pollens flows or planetary rings.

7.4. Theoretical and numerical analysis of corrosion models

The Diffusion Poisson Coupled Model [1] is a model of iron based alloy in a nuclear waste repository. It describes the growth of an oxide layer in this framework. The system is made of a Poisson equation on the electrostatic potential and convection-diffusion equations on the densities of charge carriers (electrons, ferric cations and oxygen vacancies), supplemented with coupled Robin boundary conditions. The DPCM model also takes into account the growth of the oxide host lattice and its dissolution, leading to moving boundary equations.

In [19], C. Chainais-Hillairet and I. Lacroix-Violet consider a simplified version of this model, where only two charge carriers are taken into account and where there is no evolution of the layer thickness. They prove the existence of a solution for the time-dependent simplified model.

P.-L. Colin, C.Chainais-Hillairet and I. Lacroix-Violet have performed in [16] the numerical analysis of the numerical scheme presented in [2] for the same model. The scheme is a Euler implicit in time scheme with Scharfetter-Gummel approximation of the convection-diffusion fluxes. They prove existence of a solution to the scheme, a priori estimates satisfied by the solution and convergence of the numerical scheme to a weak solution of the corrosion model.

Numerical experiments done for the simulation of the full DPCM model with moving boundaries shows the convergence in time towards a pseudo-steady-state. C. Chainais-Hillairet and T. O. Gallouët show in [37] the existence of pseudo-stationary solutions for some simplified versions of the DPCM model. They also propose a new scheme in order to compute directly this pseudo-steady-state. Numerical experiments show the efficiency of this method.

7.5. Variational modeling and analysis

Bose-Einstein condensates are a unique way to observe quantum effects at a (relatively) large scale. The fundamental states of such condensates are obtained as minimizers of a Gross-Pitaievskii functional. In [39], M. Goldman and B. Merlet consider the case of a two component Bose-Einstein condensate in the strong segregation regime (the energy favors spatial segregation of the two different Boson species). They identify two different regimes in the strong segregation and small healing length limit. In one of these regimes, the relevant limit is an interesting weighted isoperimetric problem which explains some of the numerical simulations of [63].

In [14], C. Cancès *et al.* show that the equations that are classically used for modeling the motion of two incompressible immiscible phases in a porous medium can be formally reinterpreted as the gradient flow of the free energy in a degenerated geometry closely related to the Wasserstein metric. This extends to realistic models the seminal approach [65] and the more recent one [64].

REALOPT Project-Team

7. New Results

7.1. Improving Branch-and-Price Methods

We have made progress on stabilization techniques and math-heuristics that have become essential components for Branch-and-Price methods.

Smoothing and proximal methods based on penalizing the deviation from the incumbent dual solution have become standards of the domain. Interpreting column generation as cutting plane strategies in the dual problem, we analyze in [26] the mechanisms on which stabilization relies. In particular, the link is established between smoothing and in-out separation strategies to derive generic convergence properties. For penalty function methods as well as for smoothing, we describe proposals for parameter self-adjusting schemes. Such schemes make initial parameter tuning less of an issue as corrections are made dynamically. Such adjustments also allow to adapt the parameters to the phase of the algorithm. We provide extensive test reports that validate our self-adjusting parameter scheme and highlight their performances. Our results also show that using smoothing in combination with penalty function yields a cumulative effect on convergence speed-ups.

Effects of stabilization techniques can be seen in practice. Routing and logistics applications are often viewed as intractable for exact optimization tools. Although such problems are naturally suited for a decomposition approach, branch-and-price-and-cut algorithms of the literature typically do not scale to the size of real-life instances. Some recent progress in stabilization techniques amongst other advances (such as diving heuristics, strong branching, and the combination with cutting plane approaches) generate new ambitions for column generation approach in solving approximately very large scale instances. Let us for instance point to the new benchmarks for the Capacitated Vehicle Routing Problem (CVRP) in [62]. The paper [24] illustrates this trend, showing exact results for freight transportation instances of a scale never considered before. Our column generation algorithm yields dual bounds and serves as the core procedure for a primal heuristic. The overall procedure is quite competitive in great part due to the convergence speed-ups resulting from efficient stabilization schemes. It typically provides optimal solutions as primal and dual bounds tend to be equal. The very large scale freight transportation instances (with up to 1,025 stations, 5,300 demands, and 12,651 rail cars) were submitted to us by our Russian partner Freight-One.

Math-heuristics have become an essential component in mixed integer programming (MIP) solvers. Extending generic MIP heuristics, our study in [28] outlines generic procedures to build primal solutions in the context of a Branch-and-Price approach and reports on their performance. Rounding the linear relaxation solution of the Dantzig-Wolfe reformulation, which is typically tighter than that of the original compact formulation, sometimes produces better solutions than state-of-the-art specialised heuristics as revealed by our numerical experiments. We focus on the so-called diving methods and their combination with diversification-intensification paradigms such as Limited Discrepancy Search, sub-MIPing, relaxation induced neighbourhood search, local branching, and strong branching. The dynamic generation of variables inherent to a column generation approach requires specific adaptation of heuristic paradigms. Our contribution lies in proposing simple strategies to get around these technical issues. Our numerical results on Generalized Assignment, Cutting Stock, and Vertex Coloring problems sets new benchmarks, highlighting the performance of diving heuristics as generic procedures in a column generation context.

7.2. Dual feasible functions

Dual-feasible functions have proved to be very effective for generating fast lower bounds and valid inequalities for integer linear programs with knapsack constraints. However, a significant limitation is that they are defined only for positive arguments. Extending the concept of dual-feasible function to the general domain and range R is not straightforward. In [10], we propose the first construction principles to obtain general functions with domain and range R , and we show that they lead to non-dominated maximal functions.

7.3. Allocation algorithms in Cloud platforms

In the context of service hosting in large-scale datacenters, we provide [11] a deep analysis of a cluster data trace recently released by Google and we focus on a number of questions which have not been addressed in previous studies. In particular, we describe the characteristics of job resource usage in terms of dynamics (how it varies with time), of correlation between jobs (identify daily and/or weekly patterns), and correlation inside jobs between the different resources (dependence of memory usage on CPU usage). From this analysis, we derive scalable formalizations of the allocation problem which encompass most job features. In [19], [22], we study one such model, where long-running services experience demand variations with a periodic (daily) pattern. Such services account for most of the overall CPU demand. This leads to an allocation problem where the classical Bin-Packing issue is augmented with the possibility to co-locate jobs whose peaks occur at different times of the day, which is bound to be more efficient than the usual approach that consists in over-provisioning for the maximum demand. We propose mathematical formulations, column generation approaches, and analyze their performance compared to standard packing heuristics (such as Best-Fit or First-Fit Decreasing). We show that taking periodicity of demand into account allows for a substantial improvement on machine utilization in the context of large-scale, state-of-the-art production datacenters, and that column generation allows to obtain quasi-optimal solutions in reasonable time.

7.4. Scheduling and placement for HPC

With the complexification of the architecture of HPC nodes (multicores, non uniform memory access, GPU and accelerators), a recent trend in application development is to explicitly express the computations as a task graph, and rely on a specialized middleware stack to make scheduling decisions and implement them. Traditional algorithms used in this community are dynamic heuristics, to cope with the unpredictability of execution times. In [17], [18] we analyze the performance of static and hybrid strategies, obtained by adding more static (resp. dynamic) features into dynamic (resp. static) strategies. Our conclusions are somehow unexpected in the sense that we prove that static-based strategies are very efficient, even in a context where performance estimations are not very good.

Another study [13] focuses on the memory-constrained case, where tasks may produce large data. A task can only be executed if all input and output data fit into memory, and a data can only be removed from memory after the completion of the task that uses it as an input data. Trees of such tasks arise in the multifrontal method of sparse matrix factorization. Minimizing the peak memory required on a single processor is well studied, [13] extends the problem to multiple processors, where both makespan and memory need to be minimized. We study the computational complexity of this problem and provide inapproximability results even for unit weight trees. We design a series of practical heuristics achieving different trade-offs between the minimization of peak memory usage and makespan. Some of these heuristics are able to process a tree while keeping the memory usage under a given memory limit. The different heuristics are evaluated in an extensive experimental evaluation using realistic trees.

In [20], we perform another study of static, dynamic and hybrid strategies in the context of load balancing and data placement for matrix multiplication in heterogeneous machines. Through a set of extensive simulations, we analyze the behavior of static, dynamic, and hybrid strategies, and we assess the possible benefits of introducing more static knowledge and allocation decisions in runtime libraries. In [21], we consider the purely static problem, modeled as a partitioning of a square into a set of zones of prescribed areas, while minimizing the overall size of their projections onto horizontal and vertical axes. We combine two ideas from the literature (recursive partitioning, and optimal solution structure for low number of processors) to obtain a non-rectangular recursive partitioning (NRRP), whose approximation ratio is $\frac{2}{\sqrt{3}} \simeq 1.15$, improving over the previous 1.25 ratio. Moreover, we observe on a large set of realistic platforms built from CPUs and GPUs that this proposed NRRP algorithm allows to achieve very efficient partitionings on all considered cases.

7.5. Production scheduling

Together with Shunji Tanaka, from Kyoto University, we developed Lagrangian relaxation-based methods for solving min-sum shop scheduling problems. In our studies, large scale network flow formulations of the problems are suggested together with strong Lagrangian bounds based on these formulations.

In [23], we consider the flow-shop problem on two machines with sequence-independent setup times to minimize total completion time. To cope with the size of the network, filtering procedures are developed. To solve the problem to optimality, we embed the Lagrangian bounds into two branch-and-bound algorithms. The best algorithm is able to solve all 100-jobs instances of our testbed with and without setup times, thus significantly outperforming the best algorithms in the literature, which were limited to instances with 30 and 45 jobs respectively.

In [25], we propose a new dual bound for the job-shop problem with the objective of minimizing the sum of completion costs of the operations. The bound is obtained by a Lagrangian relaxation that decomposes the problem into two types of large network flow problems: one dealing with the precedence constraints among operations of a same job, and the other one satisfying the disjunctive constraints related to the machines. Numerical experiments on the just-in-time job-shop problem show that the method is able to improve the existing lower bounds significantly.

7.6. Clustering problems

Clustering problems, and in particular partitioning problems, are widespread in combinatorial optimization. The goal is to partition a set of items in subset satisfying various constraints such as knapsack constraints, cardinality constraints, connectivity constraints, and so on. Beside the PhD thesis of Jérémy Guillot that aims to develop aggregating techniques to handles large scale instances for partitioning problems, the team also study some particular versions.

In [15] we present the application of branch-and-price approaches to the automatic version of the Software Clustering Problem. To tackle this problem, we apply the Dantzig-Wolfe decomposition to a formulation from literature. Given this, we present two Column Generation (CG) approaches to solve the linear programming relaxation of the resulting reformulation: the standard CG approach, and a new approach, which we call Staged Column Generation (SCG). Also, we propose a modification to the pricing subproblem that allows to add multiple columns at each iteration of the CG. We test our algorithms in a set of 45 instances from the literature. The proposed approaches were able to improve the literature results solving all these instances to optimality. Furthermore, the SCG approach presented a considerable performance improvement regarding computational time, number of iterations and generated columns when compared with the standard CG as the size of the instances grows.

In collaboration with researchers from University Paris 6 and Paris 13, we also study the problem of partitioning a geographical area in connected parcels. A first step of this study was to cut the area in two connected parcels while minimizing the dissimilarities inside each parcels. Such partitioning is also called a bond. It happens that in series-parallel graph, a bond correspond to a circuit in the dual graph. In [12], we give a full description of the circuit polytope on series-parallel graphs. We first show the existence of a compact extended formulation. Though not being explicit, its construction process helps us to inductively provide the description in the original space. As a consequence, using the link between bonds and circuits in planar graphs, we also describe the bond polytope on series-parallel graphs.

7.7. Tour scheduling with multi-skill heterogeneous workforce

In [14], we address a multi-activity tour scheduling problem with time varying demand. The objective is to compute a team schedule for a fixed roster of employees in order to minimize the over-coverage and the under-coverage of different parallel activity demands along a planning horizon of one week. Numerous complicating constraints are present in our problem: all employees are different and can perform several different activities during the same day-shift, lunch breaks and pauses are flexible, demand is given for 15 minutes periods. Employees have feasibility and legality rules to be satisfied, but the objective function does not account for any quality measure associated with each individual's schedule. More precisely, the problem mixes simultaneously

days-off scheduling, shift scheduling, shift assignment, activity assignment, pause and lunch break assignment. To solve this problem, we developed four methods: a compact Mixed Integer Linear Programming model, a branch-and-price like approach with a nested dynamic program to solve heuristically the subproblems, a diving heuristic and a greedy heuristic based on our subproblem solver. The computational results, based on both real cases and instances derived from real cases, demonstrate that our methods are able to provide good quality solutions in a short computing time. Our algorithms are now embedded in a commercial software, which is already in use in a mini-mart company.

7.8. Traffic routing in optical networks

In [16], we consider a multi-layer network design model arising from a real-life telecommunication application where traffic routing decisions imply the installation of expensive nodal equipment. Customer requests come in the form of bandwidth reservations for a given origin destination pair. Bandwidth demands are expressed as multiples of nominal granularities. Each request must be single-path routed. Grooming several requests on the same wavelength and multiplexing wavelengths in the same optical stream allow a more efficient use of network capacity. However, each addition or withdrawal of a request from a wavelength requires optical to electrical conversion and the use of cross-connect equipment with expensive ports of high densities. The objective is to minimize the number of required ports of the cross-connect equipment. We deal with backbone optical networks, therefore with networks with a moderate number of nodes (14 to 20) but thousands of requests. Further difficulties arise from the symmetries in wavelength assignment and traffic loading. Traditional multi-commodity network flow approaches are not suited for this problem. Instead, four alternative models relying on Dantzig-Wolfe and/or Benders' decomposition are introduced and compared. The formulations are strengthened using symmetry breaking restrictions, variable domain reduction, zero-one discretization of integer variables, and cutting planes. The resulting dual bounds are compared to the values of primal solutions obtained through hierarchical optimization and rounding procedures. For realistic size instances, our best approaches provide solutions with optimality gap of approximately 5% on average in around two hours of computing time.

7.9. Dense sphere packing

In [27], we consider the sphere packing problem in arbitrary dimension: what is the maximum fraction Δ_n of the Euclidean space \mathbb{R}_n that can be covered by unit balls with pairwise disjoint interiors?

Δ_n is known for only for some small values of n , and when n grows, we only have lower bounds. A trivial lower bound states that for every n , $\Delta_n \geq 2^{-n}$. Minkowski and Hlwaka's Theorem (1905) improves this lower bound by a factor 2: $\Delta_n \geq 2 \times 2^{-n}$. Asymptotic improvements of this bound were obtained (from Rogers, 1947 up to Ball, 1992), all of them being of the form $\Delta_n \geq cn2^{-n}$ where c is a constant.

This problem has a natural reformulation in graph theoretic terms as follows: let G denote the graph whose vertices are the points of the Euclidean space and edges are pair of vertices at distance at most 2 one from the other. The independent sets of G are the sphere packings: so, finding a maximum-density sphere packing is the same as finding a maximum-density independent set in this infinite graph. By using graph theoretic arguments only, Krivelevich et al. established that $\Delta_n \geq 0.01n2^{-n}$ for sufficiently large n .

In a recent breakthrough, Venkatesh introduced the first superlinear improvement: there are infinitely many n such that $\Delta_n \geq cn \log \log n 2^{-n}$, where c is a constant. Venkatesh's result is however non-constructive.

In this joint work with C. Bachoc and P. Moustrou, we give a constructive proof of Venkatesh's lower bound.

This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the " Investments for the future " Programme IdEx Bordeaux - CPU (ANR-10-IDEX-03-02).

SELECT Project-Team

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Serge Cohen, Erwan Le Pennec, Pascal Massart, Kevin Bleakley.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification that Cathy Maugis (INSA Toulouse) designed during her PhD thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimension. In order to circumvent this drawback, Gilles Celeux, in collaboration with Mohammed Sedki (Université Paris XI) and Cathy Maugis), proposed to sort the variables using a lasso-like penalization adapted to the Gaussian mixture model context. Using this ranking to select variables, they avoid the combinatory problem of stepwise procedures. After tests on challenging simulated and real data sets, their algorithm has shown encouraging performance. Moreover, the possibility to sort the variables with their marginal likelihoods is under study. The first results are encouraging, and this approach requires no regularization hyperparameters, and is much more rapid.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has continued research aiming to select a short list of models rather a single model. This short list is declared to be compatible with the data using a p -value derived from the Kullback-Leibler distance between the model and the empirical distribution. Furthermore, the Kullback-Leibler distances at hand are estimated through nonparametric and parametric bootstrap procedures. Different strategies are compared through numerical experiments on simulated and real data sets.

Emilie Devijver, Yannig Goude and Jean-Michel Poggi have proposed a new methodology for customer segmentation, in the context of load profiles in energy consumption. The method is based on high-dimensional regression models which perform clustering and model selection at the same time. They have focused on uncovering classes corresponding to different regression models, and compute clustering and model identification in each cluster simultaneously. They have shown the feasibility of the approach on a real data set of Irish customers.

Emilie Devijver has studied a dimension-reduction method for finite mixtures of multivariate response regression models in high-dimension. The size of the response and the number of predictors may exceed the sample size. She considers jointly predictor selection and rank reduction to obtain lower-dimensional approximations of parameter matrices. A penalty, for which the model selected by penalized likelihood satisfies an oracle inequality, is given.

The detection of change-points in a spatially or time-ordered data sequence is an important problem in many fields such as genetics and finance. Kevin Bleakley, with Gérard Biau (LSTA, Paris 6 University) and David Mason (University of Delaware), has found asymptotic distributions of statistics used to detect change-points, and developed methods to provide stopping criteria (model selection) for the number of change-points found.

6.2. Statistical learning methodology and theory

Participants: Gilles Celeux, Christine Keribin, Erwan Le Pennec, Michel Prenat, Solenne Thivin, Kevin Bleakley.

Gilles Celeux has started a collaboration with Jean-Patrick Baudry on strategies to avoid traps in the EM algorithm in mixture analysis. They analyze the effect of spurious local maximizers, and regularized algorithms to avoid such solutions. They show the link that exists between the degree of regularization and slope heuristics. Moreover, their strategy to initiate the EM algorithm, embedding the solution with K components and the starting position with $K + 1$ components to avoid suboptimal solutions, has been proved to be efficient, and is extended to a more complex framework of latent block models.

In the context of algorithms that depend on distributed computing and collaborative inference, Kevin Bleakley, with Gérard Biau (LSTA, Paris 6) and Benoît Cadre (ENS Rennes), have proposed a collaborative framework that aims to estimate the unknown mean θ of a random variable X . In the model they present, a certain number of calculation units, distributed across a communication network represented by a graph, participate in the estimation of θ by sequentially receiving independent data from X while exchanging messages via a stochastic matrix A defined over the graph. They give precise conditions on the matrix A under which the statistical precision of the individual units is comparable to that of a (gold standard) virtual centralized estimate, even though each unit does not have access to all of the data.

6.3. Reliability

Participants: Yves Auffray, Gilles Celeux, Florence Ducros, Patrick Pamphile, Jana Kalawoun.

Since June 2015, in the framework of a CIFRE convention with Nexter, Florence Ducros has commenced a thesis on the modeling of aging of vehicles, supervised by Gilles Celeux and Patrick Pamphile. This thesis should lead to designing an efficient maintenance strategy according to vehicle use profiles. It will involve the estimation of mixtures and competing risk models in a highly censored setting.

Jana Kalawoun has defended her thesis supervised by Gilles Celeux, Patrick Pamphile and Maxime Montaru (CEA) on the estimation of the battery State of Charge (SoC). For vehicles powered by an electric motor, SoC estimation is essential to guarantee vehicle autonomy, as well as safe utilization. The aim is to create a reliable SoC model to closely fit battery dynamics in embedded applications (e.g., electric vehicles). The SoC is modeled by a switching Markov state-space model. Parameters are estimated by combining the EM algorithm and particle filter methods. The model is validated using real-world electric vehicle data. This model has been proved to be highly superior to a simple state space model. The optimal number of battery modes is then identified, using model selection criteria such as AIC and BIC, which has proved to be superior to cross-validation in this particular context.

In the framework of a study on the dispatch availability of Dassault Aviation business jets, Yves Auffray and Gilles Celeux have contributed to methodology aiming to discover the root causes of reliability flaws.

6.4. Statistical analysis of genomic data

Participants: Gilles Celeux, Mélina Gallopin, Christine Keribin, Yann Vasseur.

Mélina Gallopin defended her thesis supervised by Gilles Celeux, Florence Jaffrezic and Andrea Rau (INRA, animal genetics department), This thesis was concerned with modeling and model selection in the analysis of RNA-seq data. Its highlights are the following:

- Presentation of a model selection criterion for model-based clustering of annotated gene expression data. This criterion is an ICL-like criterion taking into account annotation.
- An objective comparison of discrete and continuous modeling after transformations for RNA-seq data based on a comparison of likelihoods (possibly penalized) of the possible models.
- A block diagonal covariance selection method for high dimensional Gaussian graphical models. This non-asymptotic model selection procedure is supported by strong theoretical guarantees, based on an oracle inequality and a minimax lower bound. This work was in collaboration with Emilie Devijver.

The subject of Yann Vasseur's PhD Thesis, supervised by Gilles Celeux and Marie-Laure Martin-Magniette (INRA URGV), is the inference of a regulatory network on Transcriptions Factors (TFs), which are specific genes, of *Arabidopsis thaliana*. To that purpose, a transcriptome dataset with a similar number of TFs and statistical units is available. The first aim consists of reducing the dimension of the network to avoid high-dimensional difficulties. Representing this network with a Gaussian graphical model, the following procedure has been defined:

1. *Selection step:* choose the set of TF regulators (supports) of each TF.
2. *Classification step:* deduce co-factors groups (TFs with similar expression levels) from these supports.

Thus, the reduced network would be built on the co-factors groups. Currently, several selection methods based on Gauss-LASSO and resampling procedures have been applied to the dataset. The study of stability and parameter calibration of these methods is in progress. The TFs are clustered with the Latent Block Model in a number of co-factor groups, selected with BIC or the exact ICL criterion.

In a collaboration with Marie-Laure Martin-Magniette, Cathy Maugis and Andrea Rau, Gilles Celeux has studied gene expression obtained from high-throughput sequencing technology. The focus is on the question of clustering gene expression profiles as a means to discover groups of co-expressed genes. A Poisson mixture model is proposed, using a rigorous framework for parameter estimation as well as for the choice of the appropriate number of clusters. They illustrate co-expression analyses using this approach on two real RNA-seq datasets. A set of simulation studies also compares the performance of the proposed model with that of several related approaches developed to cluster RNA-seq and serial analysis of gene expression data. The proposed method is implemented in the open-source R package `HTSCluster`, available on CRAN. It can now be compared with Gaussian mixtures obtained after relevant data transformations.

6.5. Model based-clustering for pharmacovigilance data

Participants: Gilles Celeux, Christine Keribin, Valérie Robert.

In collaboration with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki, Gilles Celeux and Christine Keribin have started research concerning the detection of associations between drugs and adverse events in the framework of the PhD of Valerie Robert. At first, this team developed a model-based clustering inspired by latent block models, which consists of co-clustering rows and columns of two binary tables, imposing the same row ranking. This enables it to highlight subgroups of individuals sharing the same drug profile, and subgroups of adverse effects and drugs with strong interactions. Furthermore, some sufficient conditions are provided to obtain the identifiability of the model, and some results are shown for simulated data. This year, the exact ICL criterion has been extended to this double block latent model. Moreover, the possible added value of this model, compared with standard contingency table analysis, is currently under scrutiny.

SEQUEL Project-Team

7. New Results

7.1. Decision-making Under Uncertainty

7.1.1. Reinforcement Learning

Nonparametric multiple change point estimation in highly dependent time series [7]

Given a heterogeneous time-series sample, the objective is to find points in time, called change points, where the probability distribution generating the data has changed. The data are assumed to have been generated by arbitrary unknown stationary ergodic distributions. No modelling, independence or mixing assumptions are made. A novel, computationally efficient, nonparametric method is proposed, and is shown to be asymptotically consistent in this general framework. The theoretical results are complemented with experimental evaluations.

Explore no more: Improved high-probability regret bounds for non-stochastic bandits [26]

This work addresses the problem of regret minimization in non-stochastic multi-armed bandit problems, focusing on performance guarantees that hold with high probability. Such results are rather scarce in the literature since proving them requires a large deal of technical effort and significant modifications to the standard, more intuitive algorithms that come only with guarantees that hold on expectation. One of these modifications is forcing the learner to sample arms from the uniform distribution at least $\Omega(\sqrt{T})$ times over T rounds, which can adversely affect performance if many of the arms are suboptimal. While it is widely conjectured that this property is essential for proving high-probability regret bounds, we show in this paper that it is possible to achieve such strong results without this undesirable exploration component. Our result relies on a simple and intuitive loss-estimation strategy called Implicit eXploration (IX) that allows a remarkably clean analysis. To demonstrate the flexibility of our technique, we derive several improved high-probability bounds for various extensions of the standard multi-armed bandit framework. Finally, we conduct a simple experiment that illustrates the robustness of our implicit exploration technique.

First-order regret bounds for combinatorial semi-bandits [27]

We consider the problem of online combinatorial optimization under semi-bandit feedback, where a learner has to repeatedly pick actions from a combinatorial decision set in order to minimize the total losses associated with its decisions. After making each decision, the learner observes the losses associated with its action, but not other losses. For this problem, there are several learning algorithms that guarantee that the learner's expected regret grows as $O(\sqrt{T})$ with the number of rounds T . In this paper, we propose an algorithm that improves this scaling to $O(\sqrt{L * T})$, where $L * T$ is the total loss of the best action. Our algorithm is among the first to achieve such guarantees in a partial-feedback scheme, and the first one to do so in a combinatorial setting.

Random-Walk Perturbations for Online Combinatorial Optimization [4]

We study online combinatorial optimization problems where a learner is interested in minimizing its cumulative regret in the presence of switching costs. To solve such problems, we propose a version of the follow-the-perturbed-leader algorithm in which the cumulative losses are perturbed by independent symmetric random walks. In the general setting, our forecaster is shown to enjoy near-optimal guarantees on both quantities of interest, making it the best known efficient algorithm for the studied problem. In the special case of prediction with expert advice, we show that the forecaster achieves an expected regret of the optimal order $O(\sqrt{n \log N})$ where n is the time horizon and N is the number of experts, while guaranteeing that the predictions are switched at most $O(\sqrt{n \log N})$ times, in expectation.

Qualitative Multi-Armed Bandits: A Quantile-Based Approach [32]

We formalize and study the multi-armed bandit (MAB) problem in a generalized stochastic setting, in which rewards are not assumed to be numerical. Instead, rewards are measured on a qualitative scale that allows for comparison but invalidates arithmetic operations such as averaging. Correspondingly, instead of characterizing an arm in terms of the mean of the underlying distribution, we opt for using a quantile of that distribution as a representative value. We address the problem of quantile-based online learning both for the case of a finite (pure exploration) and infinite time horizon (cumulative regret minimization). For both cases, we propose suitable algorithms and analyze their properties. These properties are also illustrated by means of first experimental studies.

Predicting the outcomes of every process for which an asymptotically accurate stationary predictor exists is impossible [30]

The problem of prediction consists in forecasting the conditional distribution of the next outcome given the past. Assume that the source generating the data is such that there is a stationary predictor whose error converges to zero (in a certain sense). The question is whether there is a universal predictor for all such sources, that is, a predictor whose error goes to zero if any of the sources that have this property is chosen to generate the data. This question is answered in the negative, contrasting a number of previously established positive results concerning related but smaller sets of processes.

Improved Regret Bounds for Undiscounted Continuous Reinforcement Learning [22]

We consider the problem of undiscounted reinforcement learning in continuous state space. Regret bounds in this setting usually hold under various assumptions on the structure of the reward and transition function. Under the assumption that the rewards and transition probabilities are Lipschitz, for 1-dimensional state space a regret bound of $\tilde{O}(T^{\frac{3}{4}})$ after any T steps has been given by. Here we improve upon this result by using non-parametric kernel density estimation for estimating the transition probability distributions, and obtain regret bounds that depend on the smoothness of the transition probability distributions. In particular, under the assumption that the transition probability functions are smoothly differentiable, the regret bound is shown to be $\tilde{O}(T^{\frac{2}{3}})$ asymptotically for reinforcement learning in 1-dimensional state space. Finally, we also derive improved regret bounds for higher dimensional state space.

Maximum Entropy Semi-Supervised Inverse Reinforcement Learning [9]

A popular approach to apprenticeship learning (AL) is to formulate it as an inverse reinforcement learning (IRL) problem. The MaxEnt-IRL algorithm successfully integrates the maximum entropy principle into IRL and unlike its predecessors, it resolves the ambiguity arising from the fact that a possibly large number of policies could match the expert's behavior. In this paper, we study an AL setting in which in addition to the expert's trajectories, a number of unsupervised trajectories is available. We introduce MESSI, a novel algorithm that combines MaxEnt-IRL with principles coming from semi-supervised learning. In particular, MESSI integrates the unsupervised data into the MaxEnt-IRL framework using a pairwise penalty on trajectories. Empirical results in a highway driving and grid-world problems indicate that MESSI is able to take advantage of the unsupervised trajectories and improve the performance of MaxEnt-IRL.

Direct Policy Iteration with Demonstrations [12]

We consider the problem of learning the optimal policy of an unknown Markov decision process (MDP) when expert demonstrations are available along with interaction samples. We build on classification-based policy iteration to perform a seamless integration of interaction and expert data, thus obtaining an algorithm which can benefit from both sources of information at the same time. Furthermore, we provide a full theoretical analysis of the performance across iterations providing insights on how the algorithm works. Finally, we report an empirical evaluation of the algorithm and a comparison with the state-of-the-art algorithms.

Approximate Modified Policy Iteration and its Application to the Game of Tetris [8]

Modified policy iteration (MPI) is a dynamic programming (DP) algorithm that contains the two celebrated policy and value iteration methods. Despite its generality, MPI has not been thoroughly studied, especially its approximation form which is used when the state and/or action spaces are large or infinite. In this

paper, we propose three implementations of approximate MPI (AMPI) that are extensions of the well-known approximate DP algorithms: fitted-value iteration, fitted-Q iteration, and classification-based policy iteration. We provide error propagation analysis that unify those for approximate policy and value iteration. We develop the finite-sample analysis of these algorithms, which highlights the influence of their parameters. In the classification-based version of the algorithm (CBMPI), the analysis shows that MPI's main parameter controls the balance between the estimation error of the classifier and the overall value function approximation. We illustrate and evaluate the behavior of these new algorithms in the Mountain Car and Tetris problems. Remarkably, in Tetris, CBMPI outperforms the existing DP approaches by a large margin, and competes with the current state-of-the-art methods while using fewer samples.

7.1.2. Multi-arm Bandit Theory

Simple regret for infinitely many armed bandits [11]

We consider a stochastic bandit problem with infinitely many arms. In this setting, the learner has no chance of trying all the arms even once and has to dedicate its limited number of samples only to a certain number of arms. All previous algorithms for this setting were designed for minimizing the cumulative regret of the learner. In this paper, we propose an algorithm aiming at minimizing the simple regret. As in the cumulative regret setting of infinitely many armed bandits, the rate of the simple regret will depend on a parameter β characterizing the distribution of the near-optimal arms. We prove that depending on β , our algorithm is minimax optimal either up to a multiplicative constant or up to a $\log(n)$ factor. We also provide extensions to several important cases: when β is unknown, in a natural setting where the near-optimal arms have a small variance, and in the case of unknown time horizon.

Black-box optimization of noisy functions with unknown smoothness [20]

We study the problem of black-box optimization of a function f of any dimension, given function evaluations perturbed by noise. The function is assumed to be locally smooth around one of its global optima, but this smoothness is unknown. Our contribution is an adaptive optimization algorithm, POO or parallel optimistic optimization, that is able to deal with this setting. POO performs almost as well as the best known algorithms requiring the knowledge of the smoothness. Furthermore, POO works for a larger class of functions than what was previously considered, especially for functions that are difficult to optimize, in a very precise sense. We provide a finite-time analysis of POO's performance, which shows that its error after n evaluations is at most a factor of $\sqrt{\ln n}$ away from the error of the best known optimization algorithms using the knowledge of the smoothness.

Cheap Bandits [21]

We consider stochastic sequential learning problems where the learner can observe the average reward of several actions. Such a setting is interesting in many applications involving monitoring and surveillance, where the set of the actions to observe represent some (geographical) area. The importance of this setting is that in these applications, it is actually cheaper to observe average reward of a group of actions rather than the reward of a single action. We show that when the reward is smooth over a given graph representing the neighboring actions, we can maximize the cumulative reward of learning while minimizing the sensing cost. In this paper we propose CheapUCB, an algorithm that matches the regret guarantees of the known algorithms for this setting and at the same time guarantees a linear cost again over them. As a by-product of our analysis, we establish a (p dT) lower bound on the cumulative regret of spectral bandits for a class of graphs with effective dimension d .

Truthful Learning Mechanisms for Multi-Slot Sponsored Search Auctions with Externalities [5]

Sponsored Search Auctions (SSAs) constitute one of the most successful applications of microeconomic mechanisms. In mechanism design, auctions are usually designed to incentivize advertisers to bid their truthful valuations and, at the same time, to guarantee both the advertisers and the auctioneer a non-negative utility. Nonetheless, in sponsored search auctions, the Click-Through-Rates (CTRs) of the advertisers are often unknown to the auctioneer and thus standard truthful mechanisms cannot be directly applied and must be

paired with an effective learning algorithm for the estimation of the CTRs. This introduces the critical problem of designing a learning mechanism able to estimate the CTRs at the same time as implementing a truthful mechanism with a revenue loss as small as possible compared to the mechanism that can exploit the true CTRs. Previous work showed that, when dominant-strategy truthfulness is adopted, in single-slot auctions the problem can be solved using suitable exploration-exploitation mechanisms able to achieve a cumulative regret (on the auctioneer's revenue) of order $O(T^{(2/3)})$, where T is the number of times the auction is repeated. It is also known that, when truthfulness in expectation is adopted, a cumulative regret (over the social welfare) of order $O(T^{(1/2)})$ can be obtained. In this paper, we extend the results available in the literature to the more realistic case of multi-slot auctions. In this case, a model of the user is needed to characterize how the CTR of an ad changes as its position in the allocation changes. In particular, we adopt the cascade model, one of the most popular models for sponsored search auctions, and we prove a number of novel upper bounds and lower bounds on both auctioneer's revenue loss and social welfare w.r.t. to the Vickrey-Clarke-Groves (VCG) auction. Furthermore, we report numerical simulations investigating the accuracy of the bounds in predicting the dependency of the regret on the auction parameters.

A Relative Exponential Weighing Algorithm for Adversarial Utility-based Dueling Bandits [37]

We study the K -armed dueling bandit problem which is a variation of the classical Multi-Armed Bandit (MAB) problem in which the learner receives only relative feedback about the selected pairs of arms. We propose a new algorithm called Relative Exponential-weight algorithm for Exploration and Exploitation (REX3) to handle the adversarial utility-based formulation of this problem. This algorithm is a non-trivial extension of the Exponential-weight algorithm for Exploration and Exploitation (EXP3) algorithm. We prove a finite time expected regret upper bound of order $O(\sqrt{K \ln(K)T})$ for this algorithm and a general lower bound of order $\Omega(\sqrt{KT})$. At the end, we provide experimental results using real data from information retrieval applications.

Simultaneous Optimistic Optimization on the Noiseless BBOB Testbed [15]

We experiment the SOO (Simultaneous Optimistic Optimization) global optimizer on the BBOB testbed. We report results for both the unconstrained-budget setting and the expensive setting, as well as a comparison with the DiRect algorithm to which SOO is mostly related. Overall, SOO is shown to perform rather poorly in the highest dimensions while agreeably exhibiting interesting performance for the most difficult functions, which is to be attributed to its global nature and to the fact that its design was guided by the goal of obtaining theoretically provable performance. The greedy exploration-exploitation sampling strategy underlying SOO design is also shown to be a viable alternative for the expensive setting which gives rooms for further improvements in this direction.

7.1.3. Recommendation systems

Bandits and Recommender Systems [23]

This paper addresses the on-line recommendation problem facing new users and new items; we assume that no information is available neither about users, nor about the items. The only source of information is a set of ratings given by users to some items. By on-line, we mean that the set of users, and the set of items, and the set of ratings is evolving along time and that at any moment, the recommendation system has to select items to recommend based on the currently available information, that is basically the sequence of past events. We also mean that each user comes with her preferences which may evolve along short and longer scales of time; so we have to continuously update their preferences. When the set of ratings is the only available source of information, the traditional approach is matrix factorization. In a decision making under uncertainty setting, actions should be selected to balance exploration with exploitation; this is best modeled as a bandit problem. Matrix factors provide a latent representation of users and items. These representations may then be used as contextual information by the bandit algorithm to select items. This last point is exactly the originality of this paper: the combination of matrix factorization and bandit algorithms to solve the on-line recommendation problem. Our work is driven by considering the recommendation problem as a feedback controlled loop. This leads to interactions between the representation learning, and the recommendation policy.

Collaborative Filtering as a Multi-Armed Bandit [35]

Recommender Systems (RS) aim at suggesting to users one or several items in which they might have interest. Following the feedback they receive from the user, these systems have to adapt their model in order to improve future recommendations. The repetition of these steps defines the RS as a sequential process. This sequential aspect raises an exploration-exploitation dilemma, which is surprisingly rarely taken into account for RS without contextual information. In this paper we present an explore-exploit collaborative filtering RS, based on Matrix Factorization and Bandits algorithms. Using experiments on artificial and real datasets, we show the importance and practicability of using sequential approaches to perform recommendation. We also study the impact of the model update on both the quality and the computation time of the recommendation procedure.

AUC Optimisation and Collaborative Filtering [39]

In recommendation systems, one is interested in the ranking of the predicted items as opposed to other losses such as the mean squared error. Although a variety of ways to evaluate rankings exist in the literature, here we focus on the Area Under the ROC Curve (AUC) as it is widely used and has a strong theoretical underpinning. In practical recommendation, only items at the top of the ranked list are presented to the users. With this in mind, we propose a class of objective functions over matrix factorisations which primarily represent a smooth surrogate for the real AUC, and in a special case we show how to prioritise the top of the list. The objectives are differentiable and optimised through a carefully designed stochastic gradient-descent-based algorithm which scales linearly with the size of the data. In the special case of square loss we show how to improve computational complexity by leveraging previously computed measures. To understand theoretically the underlying matrix factorisation approaches we study both the consistency of the loss functions with respect to AUC, and generalisation using Rademacher theory. The resulting generalisation analysis gives strong motivation for the optimisation under study. Finally, we provide computation results as to the efficacy of the proposed method using synthetic and real data.

Collaborative Filtering with Localised Ranking [16]

In recommendation systems, one is interested in the ranking of the predicted items as opposed to other losses such as the mean squared error. Although a variety of ways to evaluate rankings exist in the literature, here we focus on the Area Under the ROC Curve (AUC) as it is widely used and has a strong theoretical underpinning. In practical recommendation, only items at the top of the ranked list are presented to the users. With this in mind we propose a class of objective functions which primarily represent a smooth surrogate for the real AUC, and in a special case we show how to prioritise the top of the list. This loss is differentiable and is optimised through a carefully designed stochastic gradient-descent-based algorithm which scales linearly with the size of the data. We mitigate sample bias present in the data by sampling observations according to a certain power-law based distribution. In addition, we provide computation results as to the efficacy of the proposed method using synthetic and real data.

Collaborative Filtering with Stacked Denoising AutoEncoders and Sparse Inputs [36]

Neural networks have not been widely studied in Collaborative Filtering. For instance, no paper using neural networks was published during the Net-flix Prize apart from Salakhutdinov et al's work on Restricted Boltzmann Machine (RBM) [14]. While deep learning has tremendous success in image and speech recognition, sparse inputs received less attention and remains a challenging problem for neural networks. Nonetheless, sparse inputs are critical for collaborative filtering. In this paper, we introduce a neural network architecture which computes a non-linear matrix factorization from sparse rating inputs. We show experimentally on the movieLens and jester dataset that our method performs as well as the best collaborative filtering algorithms. We provide an implementation of the algorithm as a reusable plugin for Torch [4], a popular neural network framework.

7.1.4. Nonparametric statistics of time series***The Replacement Bootstrap for Dependent Data [31]***

Applications that deal with time-series data often require evaluating complex statistics for which each time series is essentially one data point. When only a few time series are available, bootstrap methods are used to generate additional samples that can be used to evaluate empirically the statistic of interest. In this work a novel bootstrap method is proposed, which is shown to have some asymptotic consistency guarantees under the only assumption that the time series are stationary and ergodic. This contrasts previously available results that impose mixing or finite-memory assumptions on the data. Empirical evaluation on simulated and real data, using a practically relevant and complex extrema statistic is provided.

7.1.5. Imitation and Inverse Reinforcement Learning

Inverse Reinforcement Learning in Relational Domains [24]

In this work, we introduce the first approach to the Inverse Reinforcement Learning (IRL) problem in relational domains. IRL has been used to recover a more compact representation of the expert policy leading to better generalization performances among different contexts. On the other hand, relational learning allows representing problems with a varying number of objects (potentially infinite), thus provides more generalizable representations of problems and skills. We show how these different formalisms allow one to create a new IRL algorithm for relational domains that can recover with great efficiency rewards from expert data that have strong generalization and transfer properties. We evaluate our algorithm in representative tasks and study the impact of diverse experimental conditions such as : the number of demonstrations, knowledge about the dynamics, transfer among varying dimensions of a problem, and changing dynamics.

Imitation Learning Applied to Embodied Conversational Agents [29]

Embodied Conversational Agents (ECAs) are emerging as a key component to allow human interact with machines. Applications are numerous and ECAs can reduce the aversion to interact with a machine by providing user-friendly interfaces. Yet, ECAs are still unable to produce social signals appropriately during their interaction with humans, which tends to make the interaction less instinctive. Especially, very little attention has been paid to the use of laughter in human-avatar interactions despite the crucial role played by laughter in human-human interaction. In this paper, methods for predicting when and how to laugh during an interaction for an ECA are proposed. Different Imitation Learning (also known as Apprenticeship Learning) algorithms are used in this purpose and a regularized classification algorithm is shown to produce good behavior on real data.

7.1.6. Stochastic Games

Optimism in Active Learning [3]

Active learning is the problem of interactively constructing the training set used in classification in order to reduce its size. It would ideally successively add the instance-label pair that decreases the classification error most. However, the effect of the addition of a pair is not known in advance. It can still be estimated with the pairs already in the training set. The online minimization of the classification error involves a tradeoff between exploration and exploitation. This is a common problem in machine learning for which multiarmed bandit, using the approach of Optimism in the Face of Uncertainty, has proven very efficient these last years. This paper introduces three algorithms for the active learning problem in classification using Optimism in the Face of Uncertainty. Experiments lead on built-in problems and real world datasets demonstrate that they compare positively to state-of-the-art methods.

Bayesian Credible Intervals for Online and Active Learning of Classification Trees [13]

Classification trees have been extensively studied for decades. In the online learning scenario, a whole class of algorithms for decision trees has been introduced, called incremental decision trees. In the case where subtrees may not be discarded, an incremental decision tree can be seen as a sequential decision process, consisting in deciding to extend the existing tree or not. This problem involves an trade-off between exploration and exploitation, which is addressed in recent work with the use of Hoeffding's bounds. This paper proposes to use Bayesian Credible Intervals instead, in order to get the most out of the knowledge of the output's

distribution's shape. It also studies the case of Active Learning in such a tree following the Optimism in the Face of Uncertainty paradigm. Two novel algorithms are introduced for the online and active learning problems. Evaluations on real-world datasets show that these algorithms compare positively to state-of-the-art.

Optimism in Active Learning with Gaussian Processes [14]

In the context of Active Learning for classification, the classification error depends on the joint distribution of samples and their labels which is initially unknown. The minimization of this error requires estimating this distribution. Online estimation of this distribution involves a trade-off between exploration and exploitation. This is a common problem in machine learning for which multi-armed bandit theory, building upon Optimism in the Face of Uncertainty, has been proven very efficient these last years. We introduce two novel algorithms that use Optimism in the Face of Uncertainty along with Gaussian Processes for the Active Learning problem. The evaluation lead on real world datasets shows that these new algorithms compare positively to state-of-the-art methods.

Approximate Dynamic Programming for Two-Player Zero-Sum Markov Games [28]

This paper provides an analysis of error propagation in Approximate Dynamic Programming applied to zero-sum two-player Stochastic Games. We provide a novel and unified error propagation analysis in L_p -norm of three well-known algorithms adapted to Stochastic Games (namely Approximate Value Iteration, Approximate Policy Iteration and Approximate Generalized Policy Iteration). We show that we can achieve a stationary policy which is $2\gamma + (1-\gamma)$ 2-optimal, where ϵ is the value function approximation error and δ is the approximate greedy operator error. In addition, we provide a practical algorithm (AGPI-Q) to solve infinite horizon γ -discounted two-player zero-sum Stochastic Games in a batch setting. It is an extension of the Fitted-Q algorithm (which solves Markov Decision Processes from data) and can be non-parametric. Finally, we demonstrate experimentally the performance of AGPI-Q on a simultaneous two-player game, namely Alesia.

7.2. Statistical analysis of time series

7.2.1. Automata Learning

Non-negative Spectral Learning for Linear Sequential Systems [18]

Method of moments (MoM) has recently become an appealing alternative to standard iterative approaches like Expectation Maximization (EM) to learn latent variable models. In addition, MoM-based algorithms come with global convergence guarantees in the form of finite sample bounds. However, given enough computation time, by using restarts and heuristics to avoid local optima, iterative approaches often achieve better performance. We believe that this performance gap is in part due to the fact that MoM-based algorithms can output negative probabilities. By constraining the search space, we propose a non-negative spectral algorithm (NNSpectral) avoiding computing negative probabilities by design. NNSpectral is compared to other MoM-based algorithms and EM on synthetic problems of the PAutomaC challenge. Not only, NNSpectral outperforms other MoM-based algorithms, but also, achieves very competitive results in comparison to EM.

Learning of scanning strategies for electronic support using predictive state representations [17]

In Electronic Support, a receiver must monitor a wide frequency spectrum in which threatening emitters operate. A common approach is to use sensors with high sensitivity but a narrow band-width. To maintain surveillance over the whole spectrum, the sensor has to sweep between frequency bands but requires a scanning strategy. Search strategies are usually designed prior to the mission using an approximate knowledge of illumination patterns. This often results in open-loop policies that cannot take advantage of previous observations. As pointed out in past researches, these strategies lack of robustness to the prior. We propose a new closed loop search strategy that learns a stochastic model of each radar using predictive state representations. The learning algorithm benefits from the recent advances in spectral learning and rank minimization using nuclear norm penalization.

Spectral learning with proper probabilities for finite state automation [19]

Probabilistic Finite Automaton (PFA), Probabilistic Finite State Transducers (PFST) and Hidden Markov Models (HMM) are widely used in Automatic Speech Recognition (ASR), Text-to-Speech (TTS) systems and Part Of Speech (POS) tagging for language modeling. Traditionally, unsupervised learning of these latent variable models is done by Expectation-Maximization (EM)-like algorithms, as the Baum-Welch algorithm. In a recent alternative line of work, learning algorithms based on spectral properties of some low order moments matrices or tensors were proposed. In comparison to EM, they are orders of magnitude faster and come with theoretical convergence guarantees. However, returned models are not ensured to compute proper distributions. They often return negative values that do not sum to one, limiting their applicability and preventing them to serve as an initialization to EM-like algorithms. In this paper, we propose a new spectral algorithm able to learn a large range of models constrained to return proper distributions. We assess its performances on synthetic problems from the PAutomatC challenge and real datasets extracted from Wikipedia. Experiments show that it outperforms previous spectral approaches as well as the Baum-Welch algorithm with random restarts, in addition to serve as an efficient initialization step to EM-like algorithms.

7.3. Statistical Learning and Bayesian Analysis

7.3.1. Prediction of Sequences of Structured and Unstructured Data

Operator-valued Kernels for Learning from Functional Response Data [6]

In this paper we consider the problems of supervised classification and regression in the case where attributes and labels are functions: a data is represented by a set of functions, and the label is also a function. We focus on the use of reproducing kernel Hilbert space theory to learn from such functional data. Basic concepts and properties of kernel-based learning are extended to include the estimation of function-valued functions. In this setting, the representer theorem is restated, a set of rigorously defined infinite-dimensional operator-valued kernels that can be valuably applied when the data are functions is described, and a learning algorithm for nonlinear functional data analysis is introduced. The methodology is illustrated through speech and audio signal processing experiments.

7.4. Applications

7.4.1. Software development

An Experimental Protocol for Analyzing the Accuracy of Software Error Impact Analysis [25]

In software engineering, error impact analysis consists in predicting the software elements (e.g. modules, classes, methods) potentially impacted by a change. Impact analysis is required to optimize the testing effort. In this paper we present a new protocol to analyze the accuracy of impact analysis. This protocol uses mutation testing to simulate changes that introduce errors. To this end, we introduce a variant of call graphs we name the "use graph" of a software which may be computed efficiently. We apply this protocol to two open-source projects and correctly predict the impact of 30

A Learning Algorithm for Change Impact Prediction: Experimentation on 7 Java Applications [41]

Change impact analysis consists in predicting the impact of a code change in a software application. In this paper, we take a learning perspective on change impact analysis and consider the problem formulated as follows. The artifacts that are considered are methods of object-oriented software; the change under study is a change in the code of the method, the impact is the test methods that fail because of the change that has been performed. We propose an algorithm, called LCIP that learns from past impacts to predict future impacts. To evaluate our system, we consider 7 Java software applications totaling 214,000+ lines of code. We simulate 17574 changes and their actual impact through code mutations, as done in mutation testing. We find that LCIP can predict the impact with a precision of 69

7.4.2. Spoken Dialogue Systems

Human-Machine Dialogue as a Stochastic Game [10]

In this paper, an original framework to model human-machine spoken dialogues is proposed to deal with co-adaptation between users and Spoken Dialogue Systems in non-cooperative tasks. The conversation is modeled as a Stochastic Game: both the user and the system have their own preferences but have to come up with an agreement to solve a non-cooperative task. They are jointly trained so the Dialogue Manager learns the optimal strategy against the best possible user. Results obtained by simulation show that non-trivial strategies are learned and that this framework is suitable for dialogue modeling.

SIERRA Project-Team

7. New Results

7.1. On the Global Linear Convergence of Frank-Wolfe Optimization Variants

Participant: Simon Lacoste-Julien [correspondent].

Collaboration with Martin Jaggi (ETH Zurich).

The Frank-Wolfe (FW) optimization algorithm has lately re-gained popularity thanks in particular to its ability to nicely handle the structured constraints appearing in machine learning applications. However, its convergence rate is known to be slow (sublinear) when the solution lies at the boundary. A simple less-known fix is to add the possibility to take 'away steps' during optimization, an operation that importantly does not require a feasibility oracle. In this paper [17], we highlight and clarify several variants of the Frank-Wolfe optimization algorithm that have been successfully applied in practice: away-steps FW, pairwise FW, fully-corrective FW and Wolfe's minimum norm point algorithm, and prove for the first time that they all enjoy global linear convergence, under a weaker condition than strong convexity of the objective. The constant in the convergence rate has an elegant interpretation as the product of the (classical) condition number of the function with a novel geometric quantity that plays the role of a 'condition number' of the constraint set. We provide pointers to where these algorithms have made a difference in practice, in particular with the flow polytope, the marginal polytope and the base polytope for submodular optimization.

7.2. Barrier Frank-Wolfe for Marginal Inference

Participant: Simon Lacoste-Julien [correspondent].

Collaboration with Rahul G. Krishnan [correspondent] and David Sontag (NYU).

In [16], we introduce a globally-convergent algorithm for optimizing the tree-reweighted (TRW) variational objective over the marginal polytope. The algorithm is based on the conditional gradient method (Frank-Wolfe) and moves pseudomarginals within the marginal polytope through repeated maximum a posteriori (MAP) calls. This modular structure enables us to leverage black-box MAP solvers (both exact and approximate) for variational inference, and obtains more accurate results than tree-reweighted algorithms that optimize over the local consistency relaxation. Theoretically, we bound the sub-optimality for the proposed algorithm despite the TRW objective having unbounded gradients at the boundary of the marginal polytope. Empirically, we demonstrate the increased quality of results found by tightening the relaxation over the marginal polytope as well as the spanning tree polytope on synthetic and real-world instances.

7.3. Sequential Kernel Herding: Frank-Wolfe Optimization for Particle

Filtering

Participants: Simon Lacoste-Julien [correspondent], Francis Bach.

Collaboration with Fredrik Lindsten (University of Cambridge).

Recently, the Frank-Wolfe optimization algorithm was suggested as a procedure to obtain adaptive quadrature rules for integrals of functions in a reproducing kernel Hilbert space (RKHS) with a potentially faster rate of convergence than Monte Carlo integration (and "kernel herding" was shown to be a special case of this procedure). In this paper [18], we propose to replace the random sampling step in a particle filter by Frank-Wolfe optimization. By optimizing the position of the particles, we can obtain better accuracy than random or quasi-Monte Carlo sampling. In applications where the evaluation of the emission probabilities is expensive (such as in robot localization), the additional computational cost to generate the particles through optimization can be justified. Experiments on standard synthetic examples as well as on a robot localization task indicate indeed an improvement of accuracy over random and quasi-Monte Carlo sampling.

7.4. Variance Reduced Stochastic Gradient Descent with Neighbors

Participant: Simon Lacoste-Julien [correspondent].

Collaboration with Thomas Hofmann [correspondent], Aurelien Lucchi and Brian McWilliams (ETH Zurich).

Stochastic Gradient Descent (SGD) is a workhorse in machine learning, yet its slow convergence can be a computational bottleneck. Variance reduction techniques such as SAG, SVRG and SAGA have been proposed to overcome this weakness, achieving linear convergence. However, these methods are either based on computations of full gradients at pivot points, or on keeping per data point corrections in memory. Therefore speed-ups relative to SGD may need a minimal number of epochs in order to materialize. This paper [15] investigates algorithms that can exploit neighborhood structure in the training data to share and re-use information about past stochastic gradients across data points, which offers advantages in the transient optimization phase. As a side-product we provide a unified convergence analysis for a family of variance reduction algorithms, which we call memorization algorithms. We provide experimental results supporting our theory.

7.5. Rethinking LDA: Moment Matching for Discrete ICA

Participants: Anastasia Podosinnikova [correspondent], Francis Bach, Simon Lacoste-Julien.

In [21], we consider moment matching techniques for estimation in latent Dirichlet allocation (LDA). By drawing explicit links between LDA and discrete versions of independent component analysis (ICA), we first derive a new set of cumulant-based tensors, with an improved sample complexity. Moreover, we reuse standard ICA techniques such as joint diagonalization of tensors to improve over existing methods based on the tensor power method. In an extensive set of experiments on both synthetic and real datasets, we show that our new combination of tensors and orthogonal joint diagonalization techniques outperforms existing moment matching methods.

7.6. Tensorizing Neural Networks

Participant: Anton Osokin [correspondent].

Collaboration with Alexander Novikov, Dmitry Podoprikhin and Dmitry Vetrov.

Deep neural networks currently demonstrate state-of-the-art performance in several domains. At the same time, models of this class are very demanding in terms of computational resources. In particular, a large amount of memory is required by commonly used fully-connected layers, making it hard to use the models on low-end devices and stopping the further increase of the model size. In this paper [20], we convert the dense weight matrices of the fully-connected layers to the Tensor Train format such that the number of parameters is reduced by a huge factor and at the same time the expressive power of the layer is preserved. In particular, for the Very Deep VGG networks we report the compression factor of the dense weight matrix of a fully-connected layer up to 200000 times leading to the compression factor of the whole network up to 7 times.

7.7. Context-Aware CNNs for Person Head Detection

Participant: Anton Osokin [correspondent].

Collaboration with Tuan-Hung Vu [correspondent] and Ivan Laptev from the Willow project-team.

Person detection is a key problem for many computer vision tasks. While face detection has reached maturity, detecting people under a full variation of camera view-points, human poses, lighting conditions and occlusions is still a difficult challenge. In this work [23], we focus on detecting human heads in natural scenes. Starting from the recent local R-CNN object detector, we extend it with two types of contextual cues. First, we leverage person-scene relations and propose a Global CNN model trained to predict positions and scales of heads directly from the full image. Second, we explicitly model pairwise relations among objects and train a Pairwise CNN model using a structured-output surrogate loss. The Local, Global and Pairwise models are combined into a joint CNN framework. To train and test our full model, we introduce a large dataset composed of 369,846 human heads annotated in 224,740 movie frames. We evaluate our method and demonstrate improvements of person head detection against several recent baselines in three datasets. We also show improvements of the detection speed provided by our model.

7.8. Unsupervised Learning from Narrated Instruction Videos

Participants: Jean-Baptiste Alayrac [correspondent], Simon Lacoste-Julien.

Collaboration with Piotr Bojanowski, Josef Sivic and Ivan Laptev from the Willow project-team, and Nishant Agrawal.

In [29], we address the problem of automatically learning the main steps to complete a certain task, such as changing a car tire, from a set of narrated instruction videos. The contributions of this paper are three-fold. First, we develop a new unsupervised learning approach that takes advantage of the complementary nature of the input video and the associated narration. The method solves two clustering problems, one in text and one in video, applied one after each other and linked by joint constraints to obtain a single coherent sequence of steps in both modalities. Second, we collect and annotate a new challenging dataset of real-world instruction videos from the Internet. The dataset contains about 800,000 frames for five different tasks that include complex interactions between people and objects, and are captured in a variety of indoor and outdoor settings. Third, we experimentally demonstrate that the proposed method can automatically discover, in an unsupervised manner, the main steps to achieve the task and locate the steps in the input videos.

7.9. On Pairwise Cost for Multi-Object Network Flow Tracking

Participant: Simon Lacoste-Julien.

Collaboration with Visesh Chari, Ivan Laptev [correspondent] and Josef Sivic from the Willow project-team.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the “tracking-by-detection” paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose in [13] an approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real-world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method.

7.10. Multi-utility Learning: Structured-Output Learning with Multiple Annotation-Specific Loss Functions

Participant: Anton Osokin [correspondent].

Collaboration with Roman Shapovalov, Dmitry Vetrov and Pushmeet Kohli.

Structured-output learning is a challenging problem; particularly so because of the difficulty in obtaining large datasets of fully labelled instances for training. In this paper [22], we try to overcome this difficulty by presenting a multi-utility learning framework for structured prediction that can learn from training instances with different forms of supervision. We propose a unified technique for inferring the loss functions most suitable for quantifying the consistency of solutions with the given weak annotation. We demonstrate the effectiveness of our framework on the challenging semantic image segmentation problem for which a wide variety of annotations can be used. For instance, the popular training datasets for semantic segmentation are composed of images with hard-to-generate full pixel labellings, as well as images with easy-to-obtain weak annotations, such as bounding boxes around objects, or image-level labels that specify which object categories are present in an image. Experimental evaluation shows that the use of annotation-specific loss functions dramatically improves segmentation accuracy compared to the baseline system where only one type of weak annotation is used.

7.11. Convex Optimization for Parallel Energy Minimization

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

Collaboration with Alvaro Barbero, Stefanie Jegelka and Suvrit Sra.

Energy minimization has been an intensely studied core problem in computer vision. With growing image sizes (2D and 3D), it is now highly desirable to run energy minimization algorithms in parallel. But many existing algorithms, in particular, some efficient combinatorial algorithms, are difficult to parallelize. By exploiting results from convex and submodular theory, we reformulate in [47] the quadratic energy minimization problem as a total variation denoising problem, which, when viewed geometrically, enables the use of projection and reflection based convex methods. The resulting min-cut algorithm (and code) is conceptually very simple, and solves a sequence of TV denoising problems. We perform an extensive empirical evaluation comparing state-of-the-art combinatorial algorithms and convex optimization techniques. On small problems the iterative convex methods match the combinatorial max-flow algorithms, while on larger problems they offer other flexibility and important gains: (a) their memory footprint is small; (b) their straightforward parallelizability fits multi-core platforms; (c) they can easily be warm-started; and (d) they quickly reach approximately good solutions, thereby enabling faster “inexact” solutions. A key consequence of our approach based on submodularity and convexity is that it allows to combine *any arbitrary combinatorial or convex methods as subroutines*, which allows one to obtain hybrid combinatorial and convex optimization algorithms that benefit from the strengths of both.

7.12. Active-set Methods for Submodular Optimization

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

In [46], we consider submodular optimization problems such as submodular function minimization (SFM) and quadratic problems regularized by the Lovász extension; for cut functions, this corresponds respectively to graph cuts and total variation (TV) denoising. Given a submodular function with an SFM oracle, we propose a new active-set algorithm for total variation denoising, which is more flexible than existing ones; the algorithm may be seen as a local descent algorithm over ordered partitions with explicit convergence guarantees. For functions that decompose into the sum of two functions F_1 and F_2 with efficient SFM oracles, we propose a new active-set algorithm for total variation denoising (and hence for SFM by thresholding the solution at zero). This algorithm also optimizes over ordered partitions and improves over existing ones based on TV or SFM oracles for F_1 and F_2 .

7.13. Spectral Norm Regularization of Orthonormal Representations for Graph Transduction

Participant: Francis Bach [correspondent].

Collaboration with the Indian Institute of Science, Bangalore, India.

Recent literature suggests that embedding a graph on a unit sphere leads to better generalization for graph transduction. However, the choice of optimal embedding and an efficient algorithm to compute the same remains open. In this paper [25], we show that orthonormal representations, a class of unit-sphere graph embeddings are PAC learnable. Existing PAC-based analysis do not apply as the VC dimension of the function class is infinite. We propose an alternative PAC-based bound, which do not depend on the VC dimension of the underlying function class, but is related to the famous Lovasz function. The main contribution of the paper is SPORE, a SPECTral regularized ORthonormal Embedding for graph transduction, derived from the PAC bound. SPORE is posed as a non-smooth convex function over an ellipsope. These problems are usually solved as semi-definite programs (SDPs) with time complexity $O(n^6)$. We present, Infeasible Inexact proximal (IIP): an Inexact proximal method which performs subgradient procedure on an approximate projection, not necessarily feasible. IIP is more scalable than SDP, has an $O(1/\sqrt{T})$ convergence, and is generally applicable whenever a suitable approximate projection is available. We use IIP to compute SPORE where the approximate projection step is computed by FISTA, an accelerated gradient descent procedure. We show that the method has a convergence rate of $O(1/\sqrt{T})$. The proposed algorithm easily scales to 1000's of vertices, while the standard SDP computation does not scale beyond few hundred vertices. Furthermore, the analysis presented here easily extends to the multiple graph setting.

7.14. On the Equivalence between Quadrature Rules and Random Features

Participant: Francis Bach [correspondent].

In [31], we show that kernel-based quadrature rules for computing integrals can be seen as a special case of random feature expansions for positive definite kernels, for a particular decomposition that always exists for such kernels. We provide a theoretical analysis of the number of required samples for a given approximation error, leading to both upper and lower bounds that are based solely on the eigenvalues of the associated integral operator and match up to logarithmic terms. In particular, we show that the upper bound may be obtained from independent and identically distributed samples from a specific non-uniform distribution, while the lower bound is valid for any set of points. Applying our results to kernel-based quadrature, while our results are fairly general, we recover known upper and lower bounds for the special cases of Sobolev spaces. Moreover, our results extend to the more general problem of full function approximations (beyond simply computing an integral), with results in L_2 - and L_{∞} -norm that match known results for special cases. Applying our results to random features, we show an improvement of the number of random features needed to preserve the generalization guarantees for learning with Lipschitz-continuous losses.

7.15. Preconditioning of a Generalized Forward-Backward Splitting and Application to Optimization on Graphs

Participant: Loïc Landrieu [correspondent].

Collaboration with Hugo Raguet.

In [41], we present a preconditioning of a generalized forward-backward splitting algorithm for finding a zero of a sum of maximally monotone operators $\sum_{i=1}^n A_i + B$ with B cocoercive, involving only the computation of B and of the resolvent of each A_i separately. This allows in particular to minimize functionals of the form $\sum_{i=1}^n g_i + f$ with f smooth, using only the gradient of f and the proximity operator of each g_i separately. By adapting the underlying metric, such preconditioning can serve two practical purposes: first, it might accelerate the convergence, or second, it might simplify the computation of the resolvent of A_i for some i . In addition, in many cases of interest, our preconditioning strategy allows the economy of storage and computation concerning some auxiliary variables. In particular, we show how this approach can handle large-scale, non-smooth, convex optimization problems structured on graphs, which arises in many image processing or learning applications, and that it compares favourably to alternatives in the literature.

7.16. A Riemannian Low-Rank Method for Optimization over Semidefinite Matrices with Block-Diagonal Constraints

Participant: Nicolas Boumal [correspondent].

In [34], we propose a new algorithm to solve optimization problems of the form $\min f(X)$ for a smooth function f under the constraints that X is positive semidefinite and the diagonal blocks of X are small identity matrices. Such problems often arise as the result of relaxing a rank constraint (lifting). In particular, many estimation tasks involving phases, rotations, orthonormal bases or permutations fit in this framework, and so do certain relaxations of combinatorial problems such as Max-Cut. The proposed algorithm exploits the facts that (1) such formulations admit low-rank solutions, and (2) their rank-restricted versions are smooth optimization problems on a Riemannian manifold. Combining insights from both the Riemannian and the convex geometries of the problem, we characterize when second-order critical points of the smooth problem reveal KKT points of the semidefinite problem. We compare against state of the art, mature software and find that, on certain interesting problem instances, what we call the staircase method is orders of magnitude faster, is more accurate and scales better. Code is available.

7.17. Tightness of the Maximum Likelihood Semidefinite Relaxation for Angular Synchronization

Participant: Nicolas Boumal [correspondent].

Collaboration with Afonso S. Bandeira and Amit Singer.

Many maximum likelihood estimation problems are, in general, intractable optimization problems. As a result, it is common to approximate the maximum likelihood estimator (MLE) using convex relaxations. Semidefinite relaxations are among the most popular. Sometimes, the relaxations turn out to be tight. In this paper [33], we study such a phenomenon. The angular synchronization problem consists in estimating a collection of n phases, given noisy measurements of some of the pairwise relative phases. The MLE for the angular synchronization problem is the solution of a (hard) non-bipartite Grothendieck problem over the complex numbers. It is known that its semidefinite relaxation enjoys worst-case approximation guarantees. In this paper, we consider a stochastic model on the input of that semidefinite relaxation. We assume there is a planted signal (corresponding to a ground truth set of phases) and the measurements are corrupted with random noise. Even though the MLE does not coincide with the planted signal, we show that the relaxation is, with high probability, tight. This holds even for high levels of noise. This analysis explains, for the interesting case of angular synchronization, a phenomenon which has been observed without explanation in many other settings. Namely, the fact that even when exact recovery of the ground truth is impossible, semidefinite relaxations for the MLE tend to be tight (in favorable noise regimes).

7.18. Coherent Diffractive Imaging Using Randomly Coded Masks

Participant: Alexandre d'Aspremont [correspondent].

Collaboration with Matthew H. Seaberg and Joshua J. Turner.

Coherent diffractive imaging (CDI) provides new opportunities for high resolution X-ray imaging with simultaneous amplitude and phase contrast. Extensions to CDI broaden the scope of the technique for use in a wide variety of experimental geometries and physical systems. Here [44], we experimentally demonstrate a new extension to CDI that encodes additional information through the use of a series of randomly coded masks. The information gained from the few additional diffraction measurements removes the need for typical object-domain constraints; the algorithm uses prior information about the masks instead. The experiment is performed using a laser diode at 532.2 nm, enabling rapid prototyping for future X-ray synchrotron and even free electron laser experiments. Diffraction patterns are collected with up to 15 different masks placed between a CCD detector and a single sample. Phase retrieval is performed using a convex relaxation routine known as “PhaseCut” followed by a variation on Fienup’s input-output algorithm. The reconstruction quality is judged via calculation of phase retrieval transfer functions as well as by an object-space comparison between reconstructions and a lens-based image of the sample. The results of this analysis indicate that with enough masks (in this case 3 or 4) the diffraction phases converge reliably, implying stability and uniqueness of the retrieved solution.

7.19. Renegar’s Condition Number and Compressed Sensing Performance

Participants: Vincent Roulet, Nicolas Boumal, Alexandre d'Aspremont [correspondent].

Renegar’s condition number is a data-driven computational complexity measure for convex programs, generalizing classical condition numbers in linear systems. In [42], we provide evidence that for a broad class of compressed sensing problems, the worst case value of this algorithmic complexity measure taken over all signals matches the restricted eigenvalue of the observation matrix, which controls compressed sensing performance. This means that, in these problems, a single parameter directly controls computational complexity and recovery performance.

7.20. Supervised Clustering in the Data Cube

Participants: Vincent Roulet [correspondent], Fajwel Fogel, Alexandre d'Aspremont, Francis Bach.

In [43], we study a supervised clustering problem seeking to cluster either features, tasks or sample points using losses extracted from supervised learning problems. We formulate a unified optimization problem handling these three settings and derive algorithms whose core iteration complexity is concentrated in a k -means clustering step, which can be approximated efficiently. We test our methods on both artificial and realistic data sets extracted from movie reviews and 20NewsGroup.

7.21. Convex Relaxations for Permutation Problems

Participants: Fajwel Fogel [correspondent], Francis Bach, Alexandre d'Aspremont.

Collaboration with Rodolphe Jenatton.

Seriation seeks to reconstruct a linear order between variables using unsorted similarity information. It has direct applications in archeology and shotgun gene sequencing for example. In [4], we prove the equivalence between the seriation and the combinatorial 2-sum problem (a quadratic minimization problem over permutations) over a class of similarity matrices. The seriation problem can be solved exactly by a spectral algorithm in the noiseless case and we produce a convex relaxation for the 2-sum problem to improve the robustness of solutions in a noisy setting. This relaxation also allows us to impose additional structural constraints on the solution, to solve semi-supervised seriation problems. We present numerical experiments on archeological data, Markov chains and gene sequences.

7.22. Phase Recovery, MaxCut and Complex Semidefinite Programming

Participant: Alexandre d'Aspremont [correspondent].

Collaboration with Irène Waldspurger and Stéphane Mallat.

Phase retrieval seeks to recover a signal x from the amplitude $|Ax|$ of linear measurements. We cast the phase retrieval problem as a non-convex quadratic program over a complex phase vector and formulate a tractable relaxation (called PhaseCut) similar to the classical MaxCut semidefinite program. In [10], we solve this problem using a provably convergent block coordinate descent algorithm whose structure is similar to that of the original greedy algorithm in Gerchberg-Saxton, where each iteration is a matrix vector product. Numerical results show the performance of this approach over three different phase retrieval problems, in comparison with greedy phase retrieval algorithms and matrix completion formulations.

7.23. Choice of V for V -Fold Cross-Validation in Least-Squares

Participant: Sylvain Arlot [correspondent].

Collaboration with Matthieu Lerasle.

The paper [30] studies V -fold cross-validation for model selection in least-squares density estimation. The goal is to provide theoretical grounds for choosing V in order to minimize the least-squares loss of the selected estimator. We first prove a non-asymptotic oracle inequality for V -fold cross-validation and its bias-corrected version (V -fold penalization). In particular, this result implies that V -fold penalization is asymptotically optimal in the nonparametric case. Then, we compute the variance of V -fold cross-validation and related criteria, as well as the variance of key quantities for model selection performance. We show that these variances depend on V like $1 + 4/(V - 1)$, at least in some particular cases, suggesting that the performance increases much from $V = 2$ to $V = 5$ or 10, and then is almost constant. Overall, this can explain the common advice to take $V = 5$ —at least in our setting and when the computational power is limited—, as supported by some simulation experiments. An oracle inequality and exact formulas for the variance are also proved for Monte-Carlo cross-validation, also known as repeated cross-validation, where the parameter V is replaced by the number B of random splits of the data.

7.24. Gains and Losses are Fundamentally Different in Regret Minimization: The Sparse Case

Participant: Vianney Perchet [correspondent].

Collaboration with Joon Kwon.

In [38], we demonstrate that, in the classical non-stochastic regret minimization problem with d decisions, gains and losses to be respectively maximized or minimized are fundamentally different. Indeed, by considering the additional sparsity assumption (at each stage, at most s decisions incur a nonzero outcome), we derive optimal regret bounds of different orders. Specifically, with gains, we obtain an optimal regret guarantee after T stages of order $\sqrt{T \log s}$, so the classical dependency in the dimension is replaced by the sparsity size. With losses, we provide matching upper and lower bounds of order $\sqrt{Ts \log(d)/d}$, which is decreasing in d . Eventually, we also study the bandit setting, and obtain an upper bound of order $\sqrt{Ts \log(d/s)}$ when outcomes are losses. This bound is proven to be optimal up to the logarithmic factor $\sqrt{\log(d/s)}$.

7.25. Batched Bandit Problems

Participant: Vianney Perchet [correspondent].

Collaboration with Philippe Rigollet, Sylvain Chassang and Erik Snowberg.

Motivated by practical applications, chiefly clinical trials, we study in [39] the regret achievable for stochastic bandits under the constraint that the employed policy must split trials into a small number of batches. Our results show that a very small number of batches gives close to minimax optimal regret bounds. As a byproduct, we derive optimal policies with low switching cost for stochastic bandits.

7.26. Online Learning in Repeated Auctions

Participant: Vianney Perchet [correspondent].

Collaboration with Jonathan Weed and Philippe Rigollet.

Motivated by online advertising auctions, in [40] we consider repeated Vickrey auctions where goods of unknown value are sold sequentially and bidders only learn (potentially noisy) information about a good's value once it is purchased. We adopt an online learning approach with bandit feedback to model this problem and derive bidding strategies for two models: stochastic and adversarial. In the stochastic model, the observed values of the goods are random variables centered around the true value of the good. In this case, logarithmic regret is achievable when competing against well behaved adversaries. In the adversarial model, the goods need not be identical and we simply compare our performance against that of the best fixed bid in hindsight. We show that sublinear regret is also achievable in this case and prove matching minimax lower bounds. To our knowledge, this is the first complete set of strategies for bidders participating in auctions of this type.

SPHINX Team

7. New Results

7.1. Analysis, control and stabilization of heterogeneous systems

Motivated by the collision problem for rigid bodies in a perfect fluid, Munnier and Ramdani investigated in [9] the asymptotics of a 2D Laplace Neumann problem in a domain with cusp. The small parameter involved in the problem is the distance between the solid and the cavity's bottom. Denoting by $\alpha > 0$ the tangency exponent at the contact point, the authors prove that the solid always reaches the cavity in finite time, but with a non zero velocity for $\alpha < 2$ (real shock case), and with null velocity for $\alpha \geq 2$ (smooth landing case). The proof is based on a suitable change of variables transforming the Laplace Neumann problem into a generalized Neumann problem set on a domain containing a horizontal rectangle whose length tends to infinity as the solid approached the cavity.

The paper [14] presents the first positive result on approximate controllability for bilinear Schrödinger equations in presence of mixed spectrum when the interaction term is unbounded.

In [15], Tucsnak, Valein and Wu study the numerical approximation of the solutions of a class of abstract parabolic time optimal control problems. The main results assert that, provided that the target is a closed ball centered at the origin and of positive radius, the optimal time and the optimal controls of the approximate time optimal problems converge to the optimal time and to the optimal controls of the original problem. This is based on a nonsmooth data error estimate for abstract parabolic systems.

A vesicle is an elastic membrane containing a liquid and surrounded by another liquid. Such a vesicle can be found in nature or it can be created in laboratory. They can store and/or transport substances. Modeling vesicles is also a first step in order to study and understand the behavior of more complex cells such as red cells. Their studies are important for many applications, in particular in biological and physiological subjects. Recent papers have been devoted to both experimental studies to the modeling and finally to the mathematical analysis of the obtained models. There are many different models to describe the motion of the membrane and one can for instance optimize the shape in order to minimize the elastic energy of the membrane. Such a problem is tackled in [4] in the 2D case and in [6] in the 3D case. In [4], the optimization is done among convex domains whereas in [6], the authors consider the problem of minimizing the total mean curvature in order to understand the differences between the Helfrich energy and the Willmore energy. Up to now, these models are considered without any fluid.

In [13], San Martin, Takahashi and Tucsnak consider a class of low Reynolds number swimmers, of prolate spheroidal shape, which can be seen as simplified models of ciliated microorganisms. Within this model, the form of the swimmer does not change, the propelling mechanism consisting in tangential displacements of the material points of swimmer's boundary. They obtain the exact controllability of the prolate spheroidal swimmer and the existence of an optimal control problem (in the sense of the efficiency during a stroke). They also provide a method to compute an approximation of the efficiency by using explicit formulas for the Stokes system at the exterior of a prolate spheroid, with some particular tangential velocities at the fluid-solid interface. They analyze the sensitivity of this efficiency with respect to the eccentricity of the considered spheroid and show that for small positive eccentricity, the efficiency of a prolate spheroid is better than the efficiency of a sphere. Finally, they use numerical optimization tools to investigate the dependence of the efficiency on the number of inputs and on the eccentricity of the spheroid.

7.2. Inverse problems for heterogeneous systems

In [7], David Dos Santos Ferreira *et al.* obtain global stability estimates for a potential in a Schrödinger equation on an open bounded set in dimension $n = 3$ from the Dirichlet-to-Neumann map with partial data. This improves previous results where local stability was proved : the region under control was the penumbra

delimited by a source of light outside of the convex hull of the open set. These local estimates provided stability of log-log type corresponding to the uniqueness results in Calderón's inverse problem with partial data proved by Kenig, Sjöstrand and Uhlmann. The corresponding global estimates are proved in all dimensions higher than three. The estimates are based on the construction of solutions of the Schrödinger equation by complex geometrical optics developed in the anisotropic setting by Dos Santos Ferreira, Kenig, Salo and Uhlmann to solve the Calderón problem in certain admissible geometries.

In [20], David Dos Santos Ferreira *et al.* proved uniform L^p resolvent estimates for the stationary damped wave operator. Uniform L^p resolvent estimates for the Laplace operator on a compact smooth Riemannian manifold without boundary were first established by Shen on the Torus, then by Dos Santos Ferreira-Kenig-Salo for general compact manifolds and advanced further by Bourgain-Shao-Sogge-Yao. An alternative proof relying on the techniques of semiclassical Strichartz estimates allows to handle non-self-adjoint perturbations of the Laplacian and embeds very naturally in the semiclassical spectral analysis framework, and applies in the damped wave context.

In [10], Munnier and Ramdani considered the 2D inverse problem of recovering the positions and the velocities of slowly moving small rigid disks in a bounded cavity filled with a perfect fluid. Using an integral formulation, they first derive an asymptotic expansion of the DtN map of the problem as the diameters of the disks tend to zero. Then, combining a suitable choice of exponential type data and the DORT method (french acronym for Diagonalization of the Time Reversal Operator), a reconstruction method for the unknown positions and velocities is proposed. Let us emphasize here that this reconstruction method uses in the context of fluid-structure interaction problems a method which is usually used for waves inverse scattering (the DORT method).

In [24], Munnier and Ramdani proposed a new method to tackle a geometric inverse problem related to Calderón's inverse problem. More precisely, they proposed an explicit reconstruction formula for the cavity inverse problem using conformal mapping. This formula is derived by combining two ingredients: a new factorization result of the DtN map and the so-called generalized Polia-Szegö tensors of the cavity.

In [11], Ramdani, Tucsnak and Valein tackled a state estimation problem for an infinite dimensional system arising in population dynamics (a linear model for age-structured populations with spatial diffusion). Assume the initial state to be unknown, the considered inverse problem is to estimate asymptotically on time the state of the system from a locally distributed observation in both age and space. This is done by designing a Luenberger observer for the system, taking advantage of the particular spectral structure of the problem (the system has a finite number of unstable eigenvalues).

In [12], San Martin, Schwindt and Takahashi consider the geometrical inverse problem consisting in recovering an unknown obstacle in a viscous incompressible fluid by measurements of the Cauchy force on the exterior boundary. They deal with the case where the fluid equations are the non stationary Stokes system and using the enclosure method, they can recover the convex hull of the obstacle and the distance from a point to the obstacle. With the same method, they can obtain the same result in the case of a linear fluid-structure system composed by a rigid body and a viscous incompressible fluid. They also tackle the corresponding nonlinear systems: the Navier-Stokes system and a fluid-structure system with free boundary. Using complex spherical waves, they obtain some partial information on the distance from a point to the obstacle.

7.3. Numerical analysis and simulation of heterogeneous systems

In optics, metamaterials (also known as negative or left-handed materials), have known a growing interest in the last two decades. These artificial composite materials exhibit the property of having negative dielectric permittivity and magnetic permeability in a certain range of frequency, leading hence to materials with negative refractive index and super lens effects. In [5], Bunoiu and Ramdani studied a complex wave system involving such materials. More precisely, they consider a periodic homogenization problem involving two isotropic materials with conductivities of different signs: a classical material and a metamaterial (or negative material). Combining the \mathbf{T} -coercivity approach and the unfolding method for homogenization, they prove well-posedness results for the initial and the homogenized problems and obtain a convergence result, provided that the contrast between the two conductivities is large enough (in modulus).

Several results on domain decomposition were obtained in the frame of the collaboration of Xavier Antoine with the team of Christophe Geuzaine (Belgium). The paper [3] deals with a Schwarz-type solver for domain decomposition, the paper [8] proposes a Schwarz-type domain decomposition for high frequency electromagnetism equations, the paper [1] exposes how to use of GPESLab to solve Gross-Pitaevskii equations.

The paper [2] deals with domain decomposition for nonlinear Schrödinger equations and the book chapter [16] is focused on the modeling of Bose-Einstein condensates.

TAO Project-Team

7. New Results

7.1. Optimal Decision Making under Uncertainty

The Tao-uct-sig is working mainly on mathematical programming tools useful for power systems. In particular, we advocate a data science approach, in order to reduce the model error - which is much more critical than the optimization error, in most cases. Real data are the best way for handling uncertainties. Our main works are as follows:

- **Noisy optimization** In the context of stochastic uncertainties, noisy optimization handles the model error by simulation-based optimization. Our results include:
 - A formalization of noisy optimization in continuous domains, often poorly modeled in the evolutionary computation community [64], [6]. We also proposed heuristic rules for reaching slope $-1/2$ in log-log scale [34]. We also show that in some settings the slope -1 (classical in mathematical programming) can be recovered in evolution strategies (unpublished: <http://www.lri.fr/~teytaud/mca.pdf>), and we provided complexity bounds [20].
 - An extension of portfolio algorithms for noisy optimization. Portfolio methods are already usual in combinatorial optimization, some works exist in the continuous case, this is the first work in the noisy case [8].
 - Pragmatic approaches of noisy optimization, for improving robustness and for taking into account human expertise, including: Applying sieves methods in noisy optimization [27], paired optimization [35], and combining various policies [25].
 - Upper bounds on noisy optimization in discrete domains [5].
- **Quasi-random numbers** Continuous optimization is a key component of our works, hence we improve evolution strategies (which have simplicity and convenience qualities) by quasi-random numbers (showing that even in simple cases it is beneficial [52], and provides great improvements in highly multimodal cases (unpublished, <http://www.lri.fr/~teytaud/qrr.pdf>)). We also developed criteria and testbeds, pointing out some key points not widely studied in the optimization literature [26]. We also extended our earlier results in parallel optimization to additional algorithms [30], and used cutting planes as in the ellipsoid method, hence combining the best of both worlds, i.e. fast rates from cutting planes methodologies and parallel behavior as in evolution strategies [36].
- **Dynamical problems** The dynamical nature of power systems is critical, as transient regimes, ramping constraints are ubiquitous in unit commitment and dispatch. Optimizing policies, with their temporal components, is a challenge when the high dimension and the nonlinearities are taken into account. Games provide a nice testbed for experiments and are used in several of our works. We provided:
 - An original algorithm for learning opening books, by an unexpected use of random seeds [32]. The principle is to randomly sample policies, by modifying the random seed. This can be used for any stochastic policy: we generate thousands of deterministic policies (by setting the random seed to arbitrary values) and select the best ones. This can be applied for games (always the most convincing application for a proof of concept), and any control problem where stochastic policies are available.
 - An extension of the previous work for dynamically adapting the probability distribution for specific positions [51]. This work provides a MCTS without the scalability limitations of MCTS. This work might give birth to many future works.

7.2. Continuous Optimization

- **Markov Chain Analysis of Evolution Strategies** The theory of Markov chains with discrete time and continuous state space turns out to be very useful to analyze the convergence of adaptive evolution strategies (including simplified versions of the state-of-the-art CMA-ES). Exploiting invariance of the algorithms, we can indeed construct homogeneous Markov chains underlying the algorithms whose stability implies the linear convergence of the algorithm [65]. We have also shown how the convergence on constrained problems can be analyzed with Markov chains theory [10].

However the stability can be very difficult to prove; even the irreducibility can be very challenging to prove with current Markov chain theory. We have hence been developing new theoretical tools exploiting deterministic control models to prove more easily the irreducibility and T-chain property of general Markov chains [67]. Those theoretical tools can be applied to the optimization algorithms we are interested in, and trivialize some stability proofs [1], [10].

- **Benchmarking of continuous optimizers** We have been pursuing our effort towards improving the standards in benchmarking of continuous optimizers. We tackled the benchmarking of bi-objective problems and transferred and adapted standard benchmarking techniques from the single-objective optimization and classical derivative free optimization community to the field of EMO [28]. In addition, we have been rewriting part of the COCO platform to improve its modularity and make it less error prone and started its extension to multiobjective optimization.
- **Concentration inequalities for sampling without replacement** We studied the concentration of measure phenomenon in the case of sampling without replacement, which is directly relevant for a recent MCMC technique for large data sets, see [7] accepted to the Bernoulli journal.
- **Random projections for confident MCMC** In the paper [66] accepted at the NIPS "Bayesian Optimization Workshop", we study the benefit of replacing uniform subsampling by random projections in recent MCMC techniques for linear regression of tall datasets.
- **Automatic step size adaptation** We have derived a new, low-cost strategy for online adaptation of the step size in stochastic gradient descent and related algorithms [72]. This problem is of crucial importance in many machine learning algorithms, as current approaches often rely on exploring a grid of step sizes and performing a full optimization for each of them, a lengthy process.

7.3. Data Science

- **High Energy Physics** The success of the 2014 HiggsML challenge has created a willingness for structured collaboration from the High Energy Physics experiments. A working group has been set up and new challenges are currently explored. A yearly workshop has been decided, with a first edition at CERN 9-13 Nov. 2015, *DataScience@LHC*.

The challenge exemplifies a new machine learning task [58][56]: *learning to discover* evaluating the significance of a scientific discovery. It can be formally casted into a two-class classification problem, but with two major departures from a regular setting. 1) Discovery: labeled training examples of the positive class (signals) are not available and must be obtained from simulation. The learning machine can then address the "inverse problem" of predicting which events are signals in real data. 2) Evaluation: because the classes are enormously imbalanced and overlapping, the objective function of the classifier is a metric of a statistical test.

- **Personal Semantics** Our algorithm for inducing a taxonomy from a set of domain terms placed first in the international Taxonomy Induction task, part of the SemEval 2015 conference in Denver. Since then, we have developed a robust technique for discovering the domain vocabulary for a new topic using a directed crawler we created. We are currently creating hundreds of taxonomy for personal themes (hobbies, illnesses) that can be integrated into our Personal Semantics platform PTraces. The challenges for the coming year will be deploying and evaluating the taxonomies, and introducing newer machine learning methods, such as Latent Dirichlet Allocation, for better recognizing domain vocabularies.

- **Distributed system observation** The work on distributed system automated analysis and description [59][60], has been pursued through the continued development of the GAMA multi-agent framework <https://github.com/gama-platform/gama/wiki>. The simulation framework has been applied to the study of a new anytime reverse auctions protocol [53]. Philippe Caillou is associated to the young researcher ANR ACTEUR, coordinated by Patrick Taillandier (IDEES, Rouen university). With this project, a new BDI cognitive agent model, designed to be easy to use for non computer scientist, has been proposed [29] and applied to Rouen traffic simulation [57]. Finally, agent behavior has been extracted from human player logs to study the perception of emotive behaviors in board games [37].
- **Digital humanities** Amiqap and Cartolabe projects both start in 2016. The Cartolabe project applies machine learning techniques to determine comprehensible structures in unstructured data. The goal is to use raw textual data, and underspecified ontologies, to provide intuitive access to pertinent research activities in a large research organisation. Amiqap studies the relation between worker well-being and company performance, in collaboration with Mines ParisTech sociology department and La Fabrique de l'Industrie for research, Secafi and DARES for the data. These activities will benefit from Paola Tubaro's arrival (researcher CNRS in sociology and economy) in 2016.

7.4. Designing criteria

- **Criterion design and optimization methods for computer vision** On the topic of large-scale image segmentation with multiple object detection, targeting as an application the analysis of high-resolution multispectral satellite images covering the Earth, challenges are numerous: scalable complexity, finding good features to distinguish objects, designing shape statistics as well as an optimization method able to incorporate them. We propose a solution [42], [43] based on the construction of binary partition trees and on their optimization, whose cost is alleviated thanks to theoretical results reducing the search space. Concerning video segmentation, we have extended previous work, on the inclusion of shape growth constraints into classical MRF settings (graph cuts with globally optimal segmentation), to the case of multimodal sequences of medical 3D scans [19]. We also studied a new family of metrics in [9], together with a redefinition of the associated gradient and practical ways to compute it. This allows the consideration of new types of priors on planar curve evolution, such as piecewise-rigid motions. Surprisingly, the problem of finding the best piecewise-rigid approximation of a motion turns out to be convex, and to be linked to sparsity approaches.
- **Algorithm selection and configuration** Two PhD theses are still related to the former *Crossing the Chasm* SIG: Nacim Belkhir has worked on inline parameter tuning for the CMA-ES algorithm in the context of a large number of cores [21], and is now using surrogate models to compute the features of expensive continuous optimization (submitted). François Gonard's PhD is dedicated to algorithm selection. The original application domain is that of expensive car industry simulations (within the IRT-ROM project). Initial results concern combinatorial optimization, and François obtained a "Honorable mention from the jury" for his submission to the ICON Challenge (<http://iconchallenge.insight-centre.org/>), for its original approach coupling a pre-scheduler and an algorithm selector. A paper describing the algorithms and analyzing the results has been submitted.
- **A statistical physics perspective** In the topic of MRF design, with motivating applications in large scale inference problems like traffic congestions, we have finalized in [13] an approach based on the disordered Ising model relying on approximate solutions to the Inverse Ising problem. To this specific problem we also propose new approximate solutions, compliant with the generalized belief-propagation algorithm in the static [63] and a new l_0 regularized method based on a maximum likelihood maximization for the dynamical case [11]. In fact in [63] we have developed a method adapted to the generalized belief propagation framework, aiming at addressing directly and systematically the loop corrections without loss of scalability, offering new possibilities in the context of inference by MRF models. In parallel, a better understanding of the so-called mean-field approximation when the phase space is clustered has been derived [68] giving a direct method to solve static inverse problem

in the weak coupling limit. Apart from the method point of view, some consideration over what can be said on the data has been considered, still in the topic of MRF design. In this sense, it is shown in [69] that the reconstruction of the MRF model depends strongly on how the data are gathered, and how to remove redundant data and keep a good reconstruction.

- **Multi-objective AI Planning** This activity had almost stopped since the end of the DESCARWIN ANR project. However, a productive internship resulted in some new benchmarks in the ZenoTravel domain together with an exact solver ensuring the knowledge of the true Pareto front [48], [47].

7.5. Deep Learning and Information Theory

- **Natural Gradients for Deep Learning** Deep learning is now established as a state-of-the-art technology for performing different tasks such as image or sequence processing. Nevertheless, much of the computational burden is spent on tuning the hyper-parameters. On-going work, started during the TIMCO project, is proposing, in the framework of Riemannian gradient descents, invariant algorithms for training neural networks that effectively reduce the number of arbitrary choices, e.g., affine transformations of the activation functions or shuffling of the inputs. Moreover, the Riemannian gradient descent algorithms perform as well as the state-of-the-art optimizers for neural networks, and are even faster for training complex models. The proposed approach is based on Amari's theory of information geometry and consists in practical and well-grounded approximations for computing the Fisher metric. The scope of this framework is larger than Deep Learning and encompasses any class of statistical models.
- **Training dynamical systems online without backtracking** with application to recurrent neural networks [73]. We propose an algorithm to learn the parameters of a dynamical system in an online, memoryless setting, thus requiring no backpropagation through time, and consequently scalable, avoiding the large computational and memory cost of maintaining the full gradient of the current state with respect to the parameters. The algorithm essentially maintains, at each time, a single search direction in parameter space. The evolution of this search direction is partly stochastic and is constructed in such a way to provide, at every time, an unbiased random estimate of the gradient of the loss function with respect to the parameters.
- **Approximating Bayesian predictors thanks to Laplace's rule of succession** Laplace's "add-one" rule of succession modifies the observed frequencies in a sequence of heads and tails by adding one to the observed counts. This improves prediction by avoiding zero probabilities and corresponds to a uniform Bayesian prior on the parameter. We prove that, for any exponential family of distributions, arbitrary Bayesian predictors can be approximated by taking the average of the maximum likelihood predictor and the sequential normalized maximum likelihood predictor from information theory, which generalizes Laplace's rule. The proof heavily involves the geometry provided by the Fisher information matrix. Thus it is possible to approximate Bayesian predictors without the cost of integrating or sampling in parameter space [46].

TOSCA Project-Team

7. New Results

7.1. Probabilistic numerical methods, stochastic modelling and applications

Participants: Mireille Bossy, Nicolas Champagnat, Madalina Deaconu, Coralie Fritsch, Benoît Henry, James Inglis, Antoine Lejay, Oana-Valeria Lupascu, Sylvain Maire, Paolo Pigato, Alexandre Richard, Denis Talay, Etienne Tanré, Denis Villemonais.

7.1.1. Published works and preprints

- M. Bossy with H. Quinteros (UCHile) submitted a paper [36] on the strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs.
- M. Bossy and J.-F. Jabir (University of Valparaíso) submitted a paper [35] on the particle approximation for Lagrangian stochastic models with specular boundary condition.
- M. Bossy with N. Maizi (Mines ParisTech) and O. Pourtallier (Inria) published a book chapter [31] on game theory analysis for carbon auction market through electricity market coupling hypothesis.
- M. Bossy, O. Faugeras (Inria Sophia, EPI NEUROMATHCOMP), and D. Talay published a clarification on the well-posedness of the limit equations to the mean-field N -neuron models proposed in [58] and proven the associated propagation of chaos property. They also have completed the modeling issue in [58] by discussing the well-posedness of the stochastic differential equations which govern the behavior of the ion channels and the amount of available neurotransmitters. See [15].
- M. Bossy, N. Champagnat, S. Maire and L. Violeau worked with H. Leman (CMAP, Ecole Polytechnique) and M. Yvinec (Inria Sophia, GEOMETRICA team) on Monte Carlo methods for the linear and non-linear Poisson-Boltzmann equations [14]. These methods are based on walk on spheres algorithm, simulation of diffusion processes driven by their local time, and branching Brownian motion to deal with the nonlinear case.
- Together with M. Baar and A. Bovier (Univ. Bonn), N. Champagnat studied the adaptive dynamics of populations under the assumptions of large population, rare and small mutations [34]. In this work, the three limits are taken simultaneously, contrary to the classical approach, where the limits of large population and rare mutations are taken first, and next the limit of small mutations [59]. We therefore obtain the precise range of assumptions under which these limits can be taken, and provide explicit biological conditions for which our approximation is valid.
- N. Champagnat and C. Fritsch worked with F. Campillo (Inria Sophia-Antipolis, LEMON team) on the links between a branching process and an integro-differential equation of a growth-fragmentation-death model [37]. They proved that the two representations of the model lead to the same criteria of invasion of a population in a given environment.
- Using a new method to compute the expectation of an integral with respect to a random measure, N. Champagnat and B. Henry obtained explicit formulas for the moments of the frequency spectrum in the general branching processes known as Splitting Trees, with neutral mutations and under the infinitely-many alleles model [40]. This allows them to obtain a law of large numbers for the frequency spectrum in the limit of large time.
- N. Champagnat and P.-E. Jabin (Univ. Maryland) improved significantly the description of the functional spaces in the preprint [41], devoted to the study of strong existence and pathwise uniqueness for stochastic differential equations (SDE) with rough coefficients, typically in Sobolev spaces.

- N. Champagnat and D. Villemonais obtained criteria for existence and uniqueness of quasi-stationary distributions (QSD) and Q -processes for general absorbed Markov processes [17]. A QSD is a stationary distribution conditionally on non-absorption, and the Q -process is defined as the original Markov process conditioned to never be absorbed. The criteria ensure exponential convergence of the t -marginal of the process conditioned not to be absorbed at time t , to the QSD and also the exponential ergodicity of the Q -process.
- N. Champagnat and D. Villemonais obtained criteria for existence, uniqueness and exponential convergence in total variation to QSD for general absorbed and killed diffusion processes [43], [42]. For diffusions without killing [43], the criterion obtained is equivalent to the property that a diffusion on natural scale coming down from infinity has uniformly (w.r.t. the initial condition) bounded expectation at a fixed time t . The criteria obtained for diffusion processes with killing on $[0, \infty)$ [42] combine the last criteria and conditions on the killing time only close to 0, provided ∞ is an entrance boundary.
- N. Champagnat and D. Villemonais obtained criteria for existence, uniqueness and exponential convergence in total variation to QSD for general multi-dimensional birth and death processes in \mathbb{Z}_+^d absorbed at the boundary $\mathbb{Z}_+^d \setminus \mathbb{N}^d$ [44]. These birth and death models are motivated by population dynamics and the criteria obtained assume stronger intra-specific competition than inter-specific competition. These results are the first one for such processes, except for the particular case of branching processes, which can be studied using very specific methods.
- M. Deaconu, S. Herrmann and S. Maire introduced a new method for the simulation of the exit time and position of a δ -dimensional Brownian motion from a domain. This method is based on the connexion between the δ -dimensional Bessel process and the δ -dimensional Brownian motion thanks to an explicit Bessel hitting time distribution associated with a particular curved boundary. This allows to build a fast and accurate numerical scheme for approximating the brownian hitting time [19].
- M. Deaconu and O. Lupaşcu worked with L. Beznea (Bucharest, Romania) on the probabilistic interpretation of fragmentation phenomena. They constructed a continuous time branching process and characterized its behavior by using new potential theoretical tools [12].
- M. Deaconu, O. Lupaşcu and L. Beznea (Bucharest, Romania) started a new challenging work on the description of rupture phenomena like avalanches, by using fragmentation models. The physical properties of the model are deeply involved in this study. The first results concern a stochastic equation of fragmentation and branching processes related to avalanches [13].
- M. Deaconu, B. Dumortier and E. Vincent are working with the Venathec SAS on the acoustic control of wind farms. They constructed a new approach to control wind farms with a control model based on real-time source separation. They first designed a deterministic algorithm in order to maximize the electric production of the wind farms under the legal acoustic constraints. They showed that it is a non linear knapsack optimization problem and they proposed an efficient solution in that context using a branch and bound algorithm based on continuous relaxation. This work was published at the EWEA 2015 [30].
- In [49], B. Henry showed a central limit theorem for the population counting process of a supercritical Splitting Tree in the limit of large time. Thanks to the results of [40], he also obtained a central limit theorem for the frequency spectrum of Splitting Trees with neutral mutations and under the infinitely-many alleles model.
- S. Herrmann and E. Tanré have proposed a new very efficient algorithm to simulate the first-passage-time of a one-dimensional Brownian motion over a continuous curved boundary [23].
- J. Inglis and E. Tanré together with F. Delarue and S. Rubenthaler (Univ. Nice – Sophia Antipolis) completed their study of the mean-field convergence of a highly discontinuous particle system modeling the behavior of a spiking network of neurons [21].

- In collaboration with J. Maclaurin (Inria Sophia, EPI NEUROMATHCOMP) J. Inglis has presented a general framework to rigorously study the effect of spatio-temporal noise on traveling waves and stationary patterns. In particular the framework can incorporate versions of the stochastic neural field equation that may exhibit traveling fronts, pulses or stationary patterns. They have formulated a local SDE that describes the position of the stochastic wave up until a discontinuity time, at which point the position of the wave may jump and studied the local stability of this stochastic front and the long-time behavior of the stochastic wave [50].
- A. Lejay has continued his work on the Snapping Out Brownian motion, especially with regard to the simulation issues, with potential application to brain imaging techniques [33], [53].
- A. Lejay has continued his work on the simulation of processes with either discontinuous drift (with Arturo Kohatsu-Higa, Ritsumeikan University and Kazuhiro Yasuda, Hosei University, Japan) [52] or with discontinuous coefficients (with Lionel Lenêtre and Géraldine Pichot, EPI SAGE, Irisa) [54].
- A. Lejay has continued his work on the theory of rough paths, notably with the sensitivity aspects with Laure Coutin (Univ. Toulouse III) [47].
- In collaboration with Ivan Dimov and Jean-Michel Sellier (BAS), S. Maire developed a new Monte Carlo method, called the walk on equations, to solve linear systems of equations [22].
- In collaboration with Xuan Vu, Caroline Chaux-Moulin and Nadege Thirion-Moreau, S. Maire developed a stochastic algorithm to decompose large non-negative tensors with applications in spectroscopy [28].
- In collaboration with Martin Simon, Sylvain Maire developed a variant of the walk on spheres method to deal with diffusion equations appearing in electrical impedance tomography.
- With Giang Nguyen, Sylvain Maire worked on finite differences techniques to deal with many kinds of boundary conditions that are met during the Monte Carlo simulation of diffusions [25].
- A. Richard submitted a paper [56] on the spectral representation of L^2 -indexed increment-stationary processes. The main result states that any random field (i.e. process indexed by a multidimensional parameter of a function in L^2) with stationary increments can be written as an integral against a random measure satisfying certain properties. Applications to sample path properties of a multiparameter fractional Brownian motion are exhibited.
- D. Villemonais worked with P. Del Moral (Univ. Sydney) on the conditional ergodicity of time inhomogeneous diffusion processes [48]. They proved that, conditionally on non extinction, an elliptic time-inhomogeneous diffusion process forgets its initial distribution exponentially fast. An interacting particle scheme to numerically approximate the conditional distribution is also provided.
- D. Villemonais proved a Foster-Lyapunov type criterion which ensures the α -positive recurrence of birth and death processes. This criterion also provides a non-trivial subset of the domain of attraction for quasi-stationary distributions. Finally, this study leads to a Foster-Lyapunov type criterion which ensures the exponential ergodicity of a Fleming-Viot type particle system whose particles evolve as birth and death processes. The criterion also ensures the tightness of the sequence of empirical stationary distributions considered as a family of random measures. A numerical study of the speed of convergence of the particle system is also obtained under various settings [29].
- J. Inglis and D. Talay ended their work on mean-field limits of a stochastic particle system smoothly interacting through threshold hitting-times and applications to neural networks with dendritic component [51].

7.1.2. Other works in progress

- Together with M. Andrade (Univ. Paris 7) and R. Ferrière (ENS Paris and Univ. Arizona), N. Champagnat is working on the phenomenon of clustering in populations structured by space and traits for which local adaptation favors different trait values at different spatial locations. Two methods are used and numerically validated: a Turing instability method and a Hamilton-Jacobi approximation of the population density. This work is currently being written.

- N. Champagnat and J. Claisse (Ecole Polytechnique) are currently working on the ergodic and infinite horizon controls of discrete population dynamics with almost sure extinction in finite time. This can either correspond to control problems in favor of survival or of extinction, depending on the cost function. They have proved that these two problems are related to the QSD of the processes controlled by Markov controls. This work is currently being written.
- N. Champagnat and C. Fritsch worked with F. Campillo (Inria Sophia-Antipolis, LEMON team) on the variations of the principal eigenvalue (resp. the survival probability) of an integro-differential equation (resp. branching process) of growth-fragmentation-death models with respect to an environmental parameter. This work is currently being written.
- N. Champagnat, K. Coulibaly-Pasquier (Univ. Lorraine) and D. Villemonais are currently working on general criteria for existence, uniqueness and exponential convergence in total variation to QSD for multi-dimensional diffusions in a domain absorbed at its boundary. These results both improve and simplify the existing results and methods. This work is currently being written.
- N. Champagnat and D. Villemonais are currently working on extensions of their work [17] to general penalized processes, including time-inhomogeneous Markov processes with absorption. Their method allows to improve significantly the former results of [60], [61]. This work is currently being written.
- N. Champagnat and D. Villemonais are also working on extensions of the criteria of [17] in the form of Foster-Lyapunov criteria allowing to deal with cases where the convergence of conditional distribution to the QSD is not uniform with respect to the initial distribution. This work is currently being written.
- M. Deaconu and S. Herrmann are working on the numerical approach of the time-space Dirichlet problem.
- M. Deaconu, O. Lupaşcu and L. Beznea (Bucharest, Romania) worked on the numerical scheme for the simulation of an avalanche by using the fragmentation model. This work is currently being written.
- M. Deaconu, B. Dumortier and E. Vincent are working with the Venathec SAS on the acoustic control of wind farms. They plan to submit another article to IEEE transaction on sustainable energy soon. Currently they work on handling uncertainties in the model in order to design a stochastic algorithm.
- C. Fritsch worked with F. Campillo (Inria Sophia-Antipolis, LEMON team) and O. Ovaskainen (Univ. Helsinki) about the numerical analysis of the invasion of mutant populations in a chemostat, using branching processes and integro-differential models.
- C. Fritsch started a collaboration with B. Cloez (INRA, Montpellier) on a central limit theorem of mass-structured individual-based chemostat model.
- With P. Pigato, A. Lejay has continued his work on the estimation of parameters of skew diffusions.
- Within the ANESTOC Associate Team, R. Rebolledo (Pontificia Universidad Católica de Chile) and A. Richard initiated a work on the long-term behavior of a class of non-Markovian stochastic differential equations. These equations of Volterra type can be used to model the motion of a particle subject to friction forces in a heat bath, which could also be interesting in neuroscience for ion channels.
- A. Richard and E. Tanré are working with P. Orío (CINV, Chile) on the measurement of long-range dependence in series of neuronal spikes, and are providing a leaky integrate-and-fire model with fractional noise to include this effect. So far, we produced numerical experiments that confirm the existence of memory in our model, and A. Richard and E. Tanré now work on the convergence of the statistical estimator that measures this phenomenon.

- A. Richard, E. Tanré and S. Torres (Universidad de Valparaíso, Chile) are working on the definition of a skew fractional Brownian motion. The skew Brownian motion (sBm) is a process which is partly reflected when it reaches the horizontal line, making it a natural model for the motion of a particle crossing media with different diffusion properties. The fractional sBm is a modification of this process to incorporate long-range dependences. So far, we constructed a reflected fractional Brownian motion, and we are now investigating its approximation by a discrete-time process.
- During her internship supervised by E. Tanré and Romain Veltz (NEUROMATHCOMP team), Roberta Evangelista worked on “A stochastic model of gamma phase modulated orientation selectivity”. Neurons in primary visual cortex (V1) are known to be highly selective for stimulus orientation. Recent experimental evidence has shown that, in awake monkeys, the orientation selectivity of V1 neurons is modulated by gamma oscillations. In particular, neurons’ firing rate in response to the preferred orientation changes as a function of the gamma phase of spiking. The effect is drastically reduced for non-preferred orientations. We have introduced a stochastic model of a network of orientation-dependent excitatory and inhibitory spiking neurons. We have found conditions on the parameters such that the solutions of the mathematical model reproduce the experimental behavior.
- During his internship supervised by E. Tanré and Romain Veltz (NEUROMATHCOMP team), Quentin Cormier studies numerically and theoretically a model of spiking neuron in interaction with plasticity. The synaptic weights evolve according to biological law of plasticity. We study the existence of separable time scales. During his internship, Quentin Cormier also develop a numerical code to simulate large networks of neurons evolving according to this dynamics.
- C. Graham (Ecole Polytechnique) and D. Talay have written a large part of the second volume of their series on Mathematical Foundation of Stochastic Simulation.

7.2. Financial Mathematics

Participants: Mireille Bossy, Madalina Deaconu, Antoine Lejay, Sylvain Maire, Khaled Salhi, Denis Talay, Etienne Tanré.

7.2.1. Published works and preprints

- In collaboration with Jerome Lelong and Christophe Deluigi, Sylvain Maire built a new algorithm for the automatic integration and approximation of irregular functions [18]. This algorithm is tested numerically on the pricing of multidimensional exotic options.
- In collaboration with V. Reutenauer and C. Michel (CA-CIB), D. Talay and E. Tanré worked on a model in financial mathematics including bid-ask spread cost. They study the optimal strategy to hedge an interest rate swap that pays a fixed rate against a floating rate. They present a methodology using a stochastic gradient algorithm to optimize strategies. A paper is in revision [55].

7.2.2. Other works in progress

- K. Salhi works on partial hedging of options in an incomplete market, under constraints on the initial capital of the investor and assuming that the stock price is described by a Lévy process. In this case, perfect hedging is no more possible and we talk about partial hedging and minimization of risk. K. Salhi focuses on the Conditional Value-at-Risk minimization. He tries to give a numerical approximation to the solution in this context.
- In collaboration with J. Bion-Nadal (Ecole Polytechnique and CNRS), D. Talay pursued the study of a new calibration methodology based on dynamical risk measures and stochastic control PDEs.

ABS Project-Team

6. New Results

6.1. Modeling Interfaces and Contacts

Keywords: docking, scoring, interfaces, protein complexes, Voronoi diagrams, arrangements of balls.

6.1.1. High Resolution Crystal Structures Leverage Protein Binding Affinity Predictions

Participants: Frédéric Cazals, Simon Marillet.

In collaboration with P. Boudinot, Unité de recherche en virologie et immunologie moléculaires, INRA Jouy-en-Josas.

Predicting protein binding affinities from structural data has remained elusive, a difficulty owing to the variety of protein binding modes. Using the structure-affinity-benchmark (SAB, 144 cases with bound/unbound crystal structures and experimental affinity measurements), prediction has been undertaken either by fitting a model using a handfull of pre-defined variables, or by training a complex model from a large pool of parameters (typically hundreds). The former route unnecessarily restricts the model space, while the latter is prone to overfitting.

We design models in a third tier [20], using twelve variables describing enthalpic and entropic variations upon binding, and a model selection procedure identifying the best sparse model built from a subset of these variables. Using these models, we report three main results. First, we present models yielding a marked improvement of affinity predictions. For the whole dataset, we present a model predicting K_d within one and two orders of magnitude for 48% and 79% of cases, respectively. These statistics jump to 62% and 89% respectively, for the subset of the SAB consisting of high resolution structures. Second, we show that these performances owe to a new parameter encoding interface morphology and packing properties of interface atoms. Third, we argue that interface flexibility and prediction hardness do not correlate, and that for flexible cases, a performance matching that of the whole SAB can be achieved. Overall, our work suggests that the affinity prediction problem could be partly solved using databases of high resolution complexes whose affinity is known.

6.1.2. Dissecting Interfaces of Antibody - Antigen Complexes: from Ligand Specific Features to Binding Affinity Predictions

Participants: Frédéric Cazals, Simon Marillet.

In collaboration with: P. Boudinot, Unité de recherche en virologie et immunologie moléculaires, INRA Jouy-en-Josas; M-P. Lefranc, Univ. of Montpellier 2.

B lymphocytes recognize the antigen through their membrane immunoglobulins (IG), that can also be secreted. The diversity of IG-Ag complexes challenges our understanding in terms of binding affinity and interaction specificity.

In this work [21], we dissect the interfaces of IG-Ag complexes from high resolution crystal structures. We show that global interface statistics distinguish ligand types and that interfacial side chains play a key role in the interaction. Our analysis of the relative positions of CDR identifies a remarkably conserved pattern involving seven seams between CDR, with specific variations depending on the ligand type. Finally, we show that structural features of the interface and of the partners yield binding affinity estimates of unprecedented accuracy (median absolute error of 1.02 kcal/mol).

Our findings will be of broad interest, as understanding Ag recognition at the atomic level will help guiding design of better IG targeting Ag for therapeutic or other uses.

6.2. Modeling Macro-molecular Assemblies

Keywords: macro-molecular assembly, reconstruction by data integration, proteomics, modeling with uncertainties, curved Voronoi diagrams, topological persistence.

6.2.1. Unveiling Contacts within Macro-molecular assemblies by solving Minimum Weight Connectivity Inference Problems

Participants: Frédéric Cazals, Deepesh Agarwal.

In collaboration with C. Caillouet, and D. Coudert, from the COATI project-team (Inria - I3S (CNRS, University of Nice Sophia Antipolis)).

Consider a set of oligomers listing the subunits involved in sub-complexes of a macro-molecular assembly, obtained e.g. using native mass spectrometry or affinity purification. Given these oligomers, connectivity inference (CI) consists of finding the most plausible contacts between these subunits, and minimum connectivity inference (MCI) is the variant consisting of finding a set of contacts of smallest cardinality. MCI problems avoid speculating on the total number of contacts, but yield a subset of all contacts and do not allow exploiting a priori information on the likelihood of individual contacts. In this context, we present two novel algorithms, MILP-W and MILP-W_B [14]. The former solves the *minimum weight connectivity inference* (MWCI), an optimization problem whose criterion mixes the number of contacts and their likelihood. The latter uses the former in a bootstrap fashion, to improve the sensitivity and the specificity of solution sets.

Experiments on three systems (yeast exosome, yeast proteasome lid, human eIF3), for which reference contacts are known (crystal structure, cryo electron microscopy, cross-linking), show that our algorithms predict contacts with high specificity and sensitivity, yielding a very significant improvement over previous work, typically a twofold increase in sensitivity.

The software accompanying this paper is made available in the SBL, and should prove of ubiquitous interest whenever connectivity inference from oligomers is faced.

6.3. Modeling the Flexibility of Macro-molecules

Keywords: protein, flexibility, collective coordinate, conformational sampling dimensionality reduction.

6.3.1. Hybridizing Rapidly Growing Random Trees and Basin Hopping Yields an Improved Exploration of Energy Landscapes

Participants: Frédéric Cazals, Tom Dreyfus, Christine Roth.

In collaboration with C. Robert (IBPC / CNRS, Paris).

The number of local minima of the potential energy landscape (PEL) of molecular systems generally grows exponentially with the number of degrees of freedom, so that a crucial property of PEL exploration algorithms is their ability to identify local minima which are low lying and diverse.

In this work [22], we present a new exploration algorithm, retaining the ability of basin hopping (BH) to identify local minima, and that of *transition based rapidly exploring random trees* (T-RRT) to foster the exploration of yet unexplored regions. This ability is obtained by interleaving calls to the extension procedures of BH and T-RRT, and we show tuning the balance between these two types of calls allows the algorithm to focus on low lying regions. Computational efficiency is obtained using state-of-the-art data structures, in particular for searching approximate nearest neighbors in metric spaces.

We present results for the BLN69, a protein model whose conformational space has dimension 207 and whose PEL has been studied exhaustively. On this system, we show that the propensity of our algorithm to explore low lying regions of the landscape significantly outperforms those of BH and T-RRT.

6.4. Algorithmic Foundations

Keywords: computational geometry, Computational topology, Voronoi diagrams, α -shapes, Morse theory, graph algorithm, combinatorial optimization, statistical learning.

6.4.1. Beyond Two-sample-tests: Localizing Data Discrepancies in High-dimensional Spaces

Participants: Frédéric Cazals, Alix Lhéritier.

Comparing two sets of multivariate samples is a central problem in data analysis. From a statistical standpoint, the simplest way to perform such a comparison is to resort to a non-parametric two-sample test (TST), which checks whether the two sets can be seen as i.i.d. samples of an identical unknown distribution (the null hypothesis). If the null is rejected, one wishes to identify regions accounting for this difference. In this paper [17], we present a two-stage method providing *feedback* on this difference, based upon a combination of statistical learning (regression) and computational topology methods.

Consider two populations, each given as a point cloud in \mathbb{R}^d . In the first step, we assign a label to each set and we compute, for each sample point, a discrepancy measure based on comparing an estimate of the conditional probability distribution of the label given a position versus the global unconditional label distribution. In the second step, we study the height function defined at each point by the aforementioned estimated discrepancy. Topological persistence is used to identify persistent local minima of this height function, their *basins* defining regions of points with high discrepancy and in spatial proximity.

Experiments are reported both on synthetic and real data (satellite images and handwritten digit images), ranging in dimension from $d = 2$ to $d = 784$, illustrating the ability of our method to localize discrepancies.

On a general perspective, the ability to provide feedback downstream TST may prove of ubiquitous interest in exploratory statistics and data science.

6.4.2. A Sequential Non-parametric Two-Sample Test

Participants: Frédéric Cazals, Alix Lhéritier.

Given samples from two distributions, a nonparametric two-sample test aims at determining whether the two distributions are equal or not, based on a test statistic. This statistic may be computed on the whole dataset, or may be computed on a subset of the dataset by a function trained on its complement. We propose a third tier [19], consisting of functions exploiting a sequential framework to learn the differences while incrementally processing the data. Sequential processing naturally allows optional stopping, which makes our test the first truly sequential nonparametric two-sample test.

We show that any sequential predictor can be turned into a sequential two-sample test for which a valid p -value can be computed, yielding controlled type I error. We also show that pointwise universal predictors yield consistent tests, which can be built with a nonparametric regressor based on k -nearest neighbors in particular. We also show that mixtures and switch distributions can be used to increase power, while keeping consistency.

AIRSEA Team

7. New Results

7.1. Modeling for Oceanic and Atmospheric flows

7.1.1. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Mehdi-Pierre Daou, Laurent Debreu, Florian Lemarié, Charles Pelletier, Antoine Rousseau.

7.1.1.1. Coupling heterogeneous models in hydrodynamics

The coupling of models of different kinds is gaining more and more attention, due in particular to a need for more global modeling systems encompassing different disciplines (e.g. multi-physics) and different approaches (e.g. multi-scale, nesting). In order to develop such complex systems, it is generally more pragmatic to assemble different modeling units inside a user friendly modelling software platform rather than to develop new complex global models.

In the context of hydrodynamics, global modeling systems have to couple models of different dimensions (1D, 2D or 3D) and representing different physics (Navier-Stokes, hydrostatic Navier-Stokes, shallow water...). We have been developing coupling approaches for several years, based on so-called Schwarz algorithms. Our recent contributions address the development of absorbing boundary conditions for Navier-Stokes equations [1], and of interface conditions for coupling hydrostatic and nonhydrostatic Navier-Stokes flows [2]. In the context of our partnership with ARTELIA Group (PhD thesis of Medhi Pierre Daou), implementations of Schwarz coupling algorithms have been performed for hydrodynamics industrial codes (Mascaret, Telemac and OpenFoam), using the PALM coupling software. A first implementation has been realized in an academic test case, and a second one is presently under implementation in a much more realistic context.

7.1.1.2. Ocean-atmosphere coupling

Coupling methods routinely used in regional and global climate models do not provide the exact solution to the ocean-atmosphere problem, but an approximation of one [12]. For the last few years we have been actively working on the analysis of Schwarz waveform relaxation to apply this type of iterative coupling method to air-sea coupling [59], [60], [58]. In the context of the simulation of tropical cyclone, sensitivity tests to the coupling method have been carried out using ensemble simulations (through perturbations of the coupling frequency and initial conditions). We showed that the use of the Schwarz iterative coupling methods leads to a significantly reduced spread in the ensemble results (in terms of cyclone trajectory and intensity), thus suggesting that a source of error is removed w.r.t coupling methods en vogue in existing coupled models [61].

Motivated by this encouraging result, our activities over the last year can be divided into three topics

1. *Stability and consistency analysis of existing coupling methods:* in [12] we showed that the usual methods used in the context of ocean-atmosphere coupling are prone to splitting errors because they correspond to only one iteration of an iterative process without reaching convergence. Moreover, those methods have an additional condition for the coupling to be stable even if unconditionally stable time stepping algorithms are used.
2. *Study of physics-dynamics coupling:* during the PhD-thesis of Charles Pelletier (funded by Inria) the scope is on including the formulation of physical parameterizations in the theoretical analysis of the coupling. The first months of this Ph-D were dedicated to the study of the parameterization schemes to compute air-sea fluxes. A thorough sensitivity analysis showed that several parameters within existing schemes have no influence on the resulting fluxes. A simplified scheme retaining most the complexity of complicated parameterizations has thus been designed. This new scheme has also the advantage to be more adequate to conduct the mathematical analysis of the coupling.

3. *Design of a coupled single column model*: in order to focus on specific problems of ocean-atmosphere coupling, a work on simplified equation sets has been started. The aim is to implement a one-dimensional (in the vertical direction) coupled model with physical parameterizations representative of those used in realistic models. Thanks to this simplified coupled model the objective is to develop a benchmark suite for coupled models evaluation.

These three topics are addressed through strong collaborations between the applied mathematics and the climate community. As an illustration, the PhD-thesis of Charles Pelletier is in collaboration with the LSCE (Laboratoire des Sciences du Climat et de l'Environnement).

Moreover a PPR (*Projet à partenariat renforcé*) called SIMBAD (SIMplified Boundary Atmospheric layer moDel for ocean modeling purposes) is funded by Mercator-Ocean for the next three years (from march 2015 to march 2018). The aim of this project in collaboration with Meteo-France, Ifremer, LMD, and LOCEAN is to derive a metamodel to force high-resolution oceanic operational models for which the use of a full atmospheric model is not possible due to a prohibitive computational cost.

7.1.1.3. Data assimilation for coupled models

In the context of operational meteorology and oceanography, forecast skills heavily rely on proper combination of model prediction and available observations via data assimilation techniques. Historically, numerical weather prediction is made separately for the ocean and the atmosphere in an uncoupled way. However, in recent years, fully coupled ocean-atmosphere models are increasingly used in operational centers to improve the reliability of seasonal forecasts and tropical cyclones predictions. For coupled problems, the use of separated data assimilation schemes in each medium is not satisfactory since the result of such assimilation process is generally inconsistent across the interface, thus leading to unacceptable artefacts. Hence, there is a strong need for adapting existing data assimilation techniques to the coupled framework. As part of our ERACLIM2 contribution, R. Pellerej started a PhD on that topic late 2014. So far, three general data assimilation algorithms, based on variational data assimilation techniques, have been developed and applied to a simple coupled problem. The dynamical equations of the considered problem are coupled using an iterative Schwarz domain decomposition method. The aim is to properly take into account the coupling in the assimilation process in order to obtain a coupled solution close to the observations while satisfying the physical conditions across the air-sea interface. Preliminary results shows significant improvement compared to the usual approach on this simple system.

7.1.2. Numerical Schemes for Ocean Modelling

Participants: Eric Blayo, Laurent Debreu, Florian Lemarié.

In 2015, we worked on the stability constraints for oceanic numerical models ([13]). The idea is to carry a deep analysis of these constraints in order to propose new time stepping algorithms for ocean models. Except for vertical diffusion (and possibly the external mode and bottom drag), oceanic models usually rely on explicit time-stepping algorithms subject to Courant-Friedrichs-Lewy (CFL) stability criteria. Implicit methods could be unconditionally stable, but an algebraic system must be solved at each time step and other considerations such as accuracy and efficiency are less straightforward to achieve. Depending on the target application, the process limiting the maximum allowed time-step is generally different. In this paper, we introduce offline diagnostics to predict stability limits associated with internal gravity waves, advection, diffusion, and rotation. This suite of diagnostics is applied to a set of global, regional and coastal numerical simulations with several horizontal/vertical resolutions and different numerical models. We show that, for resolutions finer than $1/2^\circ$, models with an Eulerian vertical coordinate are generally constrained by vertical advection in a few hot spots and that numerics must be extremely robust to changes in Courant number. Based on those results, we review the stability and accuracy of existing numerical kernels in vogue in primitive equations oceanic models with a focus on advective processes and the dynamics of internal waves. We emphasize the additional value of studying the numerical kernel of oceanic models in the light of coupled space-time approaches instead of studying the time schemes independently from spatial discretizations. From this study, we suggest some guidelines for the development of temporal schemes in future generation multi-purpose oceanic models.

The increase of model resolution naturally leads to the representation of a wider energy spectrum. As a result, in recent years, the understanding of oceanic submesoscale dynamics has significantly improved. However, dissipation in submesoscale models remains dominated by numerical constraints rather than physical ones. Effective resolution is limited by the numerical dissipation range, which is a function of the model numerical filters (assuming that dispersive numerical modes are efficiently removed). In [16], we present a Baroclinic Jet test case set in a zonally reentrant channel that provides a controllable test of a model capacity at resolving submesoscale dynamics. We compare simulations from two models, ROMS and NEMO, at different mesh sizes (from 20 to 2 km). Through a spectral decomposition of kinetic energy and its budget terms, we identify the characteristics of numerical dissipation and effective resolution. It shows that numerical dissipation appears in different parts of a model, especially in spatial advection-diffusion schemes for momentum equations (KE dissipation) and tracer equations (APE dissipation) and in the time stepping algorithms. Effective resolution, defined by scale-selective dissipation, is inadequate to qualify traditional ocean models with low-order spatial and temporal filters, even at high grid resolution. High-order methods are better suited to the concept and probably unavoidable. Fourth-order filters are suited only for grid resolutions less than a few kilometers and momentum advection schemes of even higher-order may be justified. The upgrade of time stepping algorithms (from filtered Leapfrog), a cumbersome task in a model, appears critical from our results, not just as a matter of model solution quality but also of computational efficiency (extended stability range of predictor-corrector schemes). Effective resolution is also shaken by the need for non scale-selective barotropic mode filters and requires carefully addressing the issue of mode splitting errors. Possibly the most surprising result is that submesoscale energy production is largely affected by spurious diapycnal mixing (APE dissipation). This result justifies renewed efforts in reducing tracer mixing errors and poses again the question of how much vertical diffusion is at work in the real ocean.

7.1.3. *Better Parameterization of the Coastline for Ocean Models*

Participants: Eric Blayo, Eugene Kazantsev, Florian Lemarié, Pierre Marchand.

We aim at the development of finer approximations of lateral boundaries and boundary conditions for NEMO, by investigating and comparing analytical and optimal control approaches.

Regarding the analytical approach, we focused on a 2D shallow water formulation, and revisited the properties of the energy and enstrophy conserving schemes in the presence of a coastline. This led us to highlight a number of problems with the enstrophy conserving scheme (sensitivity to the choice of a slip or a noslip boundary condition, non conservation of the enstrophy, numerical instability). We also proposed a corrected scheme near the boundary for the continuity equation and new values for ghost points derived from the energy conservation in order for the energy conserving scheme to take into account a coastline with some inclination with regard to the numerical grid. We also investigated the viscous case, and proposed an implementation of slip and no slip boundary conditions for the viscous term in such a case of an inclined coastline.

These results are under comparison with the optimal control approach 7.3.2 realised for the Nemo model in a similar configuration.

7.2. Model reduction / multiscale algorithms

7.2.1. *Intrusive sensitivity analysis, reduced models*

Participants: Maëlle Nodet, Clémentine Prieur.

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called "proper orthogonal decomposition" or "principal component analysis". More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural

way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [55], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound.

When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears: a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [56]. In [54], the authors provide asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model. These results were also adapted to problems related to more general models such as Shallow-Water equations, in the context of the control of an open channel [8].

Let us come back to the output of interest. Is it possible to get better error certification when the output is specified. A work in this sense has been accepted, dealing with goal oriented uncertainties assessment [7].

A collaboration has been started with Christophe Prieur (Gipsa-Lab) on the very challenging issue of sensitivity of a controlled system to its control parameters [8].

7.2.2. *Multigrid Methods for Variational Data Assimilation.*

Participants: Laurent Debreu, François-Xavier Le Dimet, Arthur Vidard.

In order to lower the computational cost of the variational data assimilation process, we investigate the use of multigrid methods to solve the associated optimal control system. On a linear advection equation, we study the impact of the regularization term on the optimal control and the impact of discretization errors on the efficiency of the coarse grid correction step. We show that even if the optimal control problem leads to the solution of an elliptic system, numerical errors introduced by the discretization can alter the success of the multigrid methods. The view of the multigrid iteration as a preconditioner for a Krylov optimization method leads to a more robust algorithm. A scale dependent weighting of the multigrid preconditioner and the usual background error covariance matrix based preconditioner is proposed and brings significant improvements. This work is summarized in ([5]).

7.3. Dealing with uncertainties

7.3.1. *Sensitivity Analysis for Forecasting Ocean Models*

Participants: Eric Blayo, Laurent Gilquin, Céline Helbert, François-Xavier Le Dimet, Elise Arnaud, Simon Nanty, Maëlle Nodet, Clémentine Prieur, Laurence Viry, Federico Zertuche.

7.3.1.1. *Scientific context*

Forecasting geophysical systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.

7.3.1.2. Estimating sensitivity indices

A first task is to develop tools for estimated sensitivity indices. In variance-based sensitivity analysis, a classical tool is the method of Sobol' [68] which allows to compute Sobol' indices using Monte Carlo integration. One of the main drawbacks of this approach is that the estimation of Sobol' indices requires the use of several samples. For example, in a d -dimensional space, the estimation of all the first-order Sobol' indices requires $d + 1$ samples. Some interesting combinatorial results have been introduced to weaken this defect, in particular by Saltelli [66] and more recently by Owen [64] but the quantities they estimate still require $O(d)$ samples.

In a recent work [71] we introduce a new approach to estimate all first-order Sobol' indices by using only two samples based on replicated latin hypercubes and all second-order Sobol' indices by using only two samples based on replicated randomized orthogonal arrays. We establish theoretical properties of such a method for the first-order Sobol' indices and discuss the generalization to higher-order indices. As an illustration, we propose to apply this new approach to a marine ecosystem model of the Ligurian sea (northwestern Mediterranean) in order to study the relative importance of its several parameters. The calibration process of this kind of chemical simulators is well-known to be quite intricate, and a rigorous and robust — i.e. valid without strong regularity assumptions — sensitivity analysis, as the method of Sobol' provides, could be of great help. The computations are performed by using CIGRI, the middleware used on the grid of the Grenoble University High Performance Computing (HPC) center. We are also applying these estimates to calibrate integrated land use transport models. As for these models, some groups of inputs are correlated, Laurent Gilquin extended the approach based on replicated designs for the estimation of grouped Sobol' indices [6].

We can now wonder what are the asymptotic properties of these new estimators, or also of more classical ones. In [54], the authors deal with asymptotic properties of the estimators. In [52], the authors establish also a multivariate central limit theorem and non asymptotic properties.

7.3.1.3. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [53] in the case where input parameters are correlated. Clémentine Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [44]. We obtained first results [45], deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs. We then adapted various algorithms for the estimation of these new indices. These algorithms make the assumption that among the potential interactions, only few are significant. Two papers have been recently accepted [43], [46]. We also considered (see the paragraph 7.3.1) the estimation of groups Sobol' indices, with a procedure based on replicated designs. These indices provide information at the level of groups, and not at a finer level, but their interpretation is still rigorous.

Céline Helbert and Clémentine Prieur supervised the PhD thesis of Simon Nanty (funded by CEA Cadarache, and defended in October, 2015). The subject of the thesis is the analysis of uncertainties for numerical codes with temporal and spatio-temporal input variables, with application to safety and impact calculation studies. This study implied functional dependent inputs. A first step was the modeling of these inputs, and a paper has been submitted [63]. The whole methodology proposed during the PhD is under advanced revision [36].

7.3.1.4. Multy-fidelity modeling for risk analysis

Federico Zertuche's PhD concerns the modeling and prediction of a digital output from a computer code when multiple levels of fidelity of the code are available. A low-fidelity output can be obtained, for example on a coarse mesh. It is cheaper, but also much less accurate than a high-fidelity output obtained on a fine mesh. In this context, we propose new approaches to relieve some restrictive assumptions of existing methods ([57], [65]): a new estimation method of the classical cokriging model when designs are not nested and a nonparametric modeling of the relationship between low-fidelity and high-fidelity levels. The PhD takes place in the REDICE consortium and in close link with industry. The first part of the thesis was also dedicated to the development of a case study in fluid mechanics with CEA in the context of the study of a nuclear reactor.

The second part of the thesis was dedicated to the development of a new sequential approach based on a course to fine wavelets algorithm. Federico Zertuche presented his work at the annual meeting of the GDR Mascot Num in 2014 [72].

7.3.1.5. *Data assimilation and second order sensitivity analysis*

Basically, in the deterministic approach, a sensitivity analysis is the evaluation of a functional depending on the state of the system and of parameters. Therefore it is natural to introduce an adjoint model. In the framework of variational data assimilation the link between all the ingredients (observations, parameters and other inputs of the model) is done through the optimality system (O.S.), therefore a sensitivity will be estimated by deriving the O.S. leading to a second order adjoint. This is done in the paper [15] in which a full second order analysis is carried out on a model of the Black Sea.

This methodology has been applied to

- **Oil Spill.** These last years have known several disasters produced by wrecking of ships and drifting platforms with severe consequences on the physical and biological environments. In order to minimize the impact of these oil spills it is necessary to predict the evolution of oil spots. Some basic models are available and some satellites provide images on the evolution of oil spots. Clearly this topic is a combination of the two previous ones: data assimilation for pollution and assimilation of images. A theoretical framework has been developed with Dr. Tran Thu Ha (iMech).
- **Data Assimilation in Supercavitation (with iMech).** Some self propelled submarine devices can reach a high speed thanks to the phenomenon of supercavitation: an air bubble is created on the nose of the device and reduces drag forces. Some models of supercavitation already exist but are working on two applications of variational methods to supercavitation:
 - Parameter identification : the models have some parameters that can not be directly measured. From observations we retrieve the unknown parameters using a classical formalism of inverse problems.
 - Shape Optimization. The question is to determine an optimum design of the shape of the engine in order to reach a maximum speed.

7.3.2. *Optimal Control of Boundary Conditions*

Participants: Christine Kazantsev, Eugene Kazantsev.

A variational data assimilation technique is applied to the identification of the optimal boundary conditions for a simplified configuration of the NEMO model. A rectangular box model placed in mid-latitudes, and subject to the classical single or double gyre wind forcing, is studied. The model grid can be rotated on a desired angle around the center of the rectangle in order to simulate the boundary approximated by a staircase-like coastlines. The solution of the model on the grid aligned with the box borders was used as a reference solution and as artificial observational data. It is shown in [9], [10] that optimal boundary has a rather complicated geometry which is neither a staircase, nor a straight line. The boundary conditions found in the data assimilation procedure bring the solution toward the reference solution allowing to correct the influence of the rotated grid (see fig. 1).

Adjoint models, necessary to variational data assimilation, have been produced by the TAPENADE software, developed by the SCIPORT team. This software is shown to be able to produce the adjoint code that can be used in data assimilation after a memory usage optimization.

7.3.3. *Non-Parametric Estimation for Kinetic Diffusions*

Participants: Clémentine Prieur, Jose Raphael Leon Ramos.

This research is the subject of a collaboration with Venezuela and is partly funded by an ECOS Nord project.

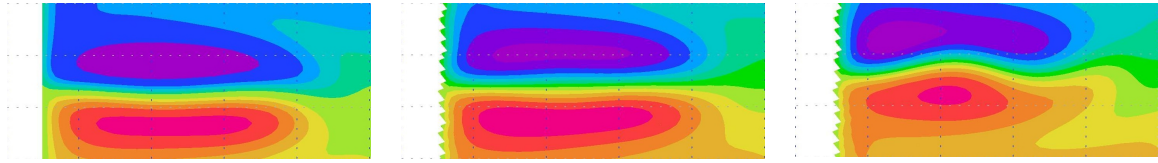


Figure 1. Sea surface elevation: reference solution on the aligned grid (left), solutions on the 30° rotated grid with optimal (center) and conventional (right) boundary conditions.

We are focusing our attention on models derived from the linear Fokker-Planck equation. From a probabilistic viewpoint, these models have received particular attention in recent years, since they are a basic example for hypercoercivity. In fact, even though completely degenerated, these models are hypoelliptic and still verify some properties of coercivity, in a broad sense of the word. Such models often appear in the fields of mechanics, finance and even biology. For such models we believe it appropriate to build statistical non-parametric estimation tools. Initial results have been obtained for the estimation of invariant density, in conditions guaranteeing its existence and unicity [40] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been accepted recently [41] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, a paper has been accepted [41], in collaboration with J.R. Leon (Caracas, Venezuela) and P. Cattiaux (Toulouse). Recursive estimators have been also proposed by the same authors in [42], also recently accepted.

Note that Professor Jose R. Leon (Caracas, Venezuela) is now funded by an international Inria Chair and will spend one year in our team, allowing to collaborate further on parameter estimation.

7.3.4. Multivariate Risk Indicators

Participants: Clémentine Prieur, Patricia Tencaliec.

Studying risks in a spatio-temporal context is a very broad field of research and one that lies at the heart of current concerns at a number of levels (hydrological risk, nuclear risk, financial risk etc.). Stochastic tools for risk analysis must be able to provide a means of determining both the intensity and probability of occurrence of damaging events such as e.g. extreme floods, earthquakes or avalanches. It is important to be able to develop effective methodologies to prevent natural hazards, including e.g. the construction of barrages.

Different risk measures have been proposed in the one-dimensional framework. The most classical ones are the return level (equivalent to the Value at Risk in finance), or the mean excess function (equivalent to the Conditional Tail Expectation CTE). However, most of the time there are multiple risk factors, whose dependence structure has to be taken into account when designing suitable risk estimators. Relatively recent regulation (such as Basel II for banks or Solvency II for insurance) has been a strong driver for the development of realistic spatio-temporal dependence models, as well as for the development of multivariate risk measurements that effectively account for these dependencies.

We refer to [47] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [67] proposed a generalization of the concept of return period in dimension greater than or equal to two. Michele et al. proposed in a recent study [48] to take into account the duration and not only the intensity of an event for designing what they call the dynamic return period. However, few studies address the issues of statistical inference in the multivariate context. In [49], [51], we proposed non parametric estimators of a multivariate extension of the CTE. As might be expected, the properties of these estimators deteriorate when considering extreme risk levels. In collaboration with Elena Di Bernardino (CNAM, Paris), Clémentine Prieur is working on the extrapolation of the above results to extreme risk levels.

Elena Di Bernardino, Véronique Maume-Deschamps (Univ. Lyon 1) and Clémentine Prieur also derived an estimator for bivariate tail [50]. The study of tail behavior is of great importance to assess risk.

With Anne-Catherine Favre (LTHE, Grenoble), Clémentine Prieur supervises the PhD thesis of Patricia Tencaliec. We are working on risk assessment, concerning flood data for the Durance drainage basin (France). The PhD thesis started in October 2013. A first paper on data reconstruction has been accepted [18]. It was a necessary step as the initial series contained many missing data.

7.4. Assimilation of Images

Participants: François-Xavier Le Dimet, Maëlle Nodet, Arthur Vidard, Nelson Feyeux, Vincent Chabot, Nicolas Papadakis.

7.4.1. Direct assimilation of image sequences

At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelops. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.

MOISE was the leader of the ANR ADDISA project dedicated to the assimilation of images, and is a member of its follow-up GeoFluids (along with EPI FLUMINANCE and CLIME, and LMD, IFREMER and Météo-France) that ended in 2013.

During the ADDISA project we developed Direct Image Sequences Assimilation (DISA) and proposed a new scheme for the regularization of optical flow problems [69], which was recently extended [17]. Thanks to the nonlinear brightness assumption, we proposed an algorithm to estimate the motion between two images, based on the minimization of a nonlinear cost function. We proved its efficiency and robustness on simulated and experimental geophysical flows [38]. As part of the ANR project GeoFluids, we are investigating new ways to define distance between a couple of images. One idea is to compare the gradient of the images rather than the actual value of the pixels. This leads to promising results. Another idea, currently under investigation, consists in comparing main structures within each image. This can be done using, for example, a wavelet representation of images. Both approaches have been compared, in particular their relative merits in dealing with observation errors, in a paper published early 2015 [4] and presented in several international conferences [21], [28].

In recent developments [11] we have also used "Level Sets" methods to describe the evolution of the images. The advantage of this approach is that it permits, thanks to the level sets function, to consider the images as a state variable of the problem. We have derived an Optimality System including the level sets of the images.

7.4.2. Optimal transport for image assimilation

Within the optimal transport project TOMMI funded by the ANR white program (started mid 2011), a new optimization scheme based on proximal splitting method has been proposed to solve the dynamic optimal transport problem. We investigate the use of optimal transport based distances for data assimilation. The study is still under investigation in the framework of N. Feyeux's PhD, but preliminary encouraging results have already been presented in [20] and an article is in preparation on this topic.

7.5. Tracking of Mesoscale Convective Systems

Participant: Clémentine Prieur.

We are interested in the tracking of mesoscale convective systems. A particular region of interest is West Africa. Data and hydrological expertise is provided by T. Vischel and T. Lebel (LTHE, Grenoble).

A first approach involves adapting the multiple hypothesis tracking (MHT) model originally designed by the NCAR (National Centre for Atmospheric Research) for tracking storms [70] to the data for West Africa. With A. Makris (working on a post-doctoral position), we proposed a Bayesian approach [62], which consists in considering that the state at time t is composed on one hand by the events (birth, death, splitting, merging) and on the other hand by the targets' attributes (positions, velocities, sizes, ...). The model decomposes the state into two sub-states: the events and the targets positions/attributes. The events are updated first and are conditioned to the previous targets sub-state. Then given the new events the target substate is updated. A simulation study allowed to verify that this approach improves the frequentist approach by Storlie et al. (2009). It has been tested on simulations [62] and investigated in the specific context of real data on West Africa [35]. Using PHD (probability hypothesis density) filters adapted to our problem, generalizing recent developments in particle filtering for spatio-temporal branching processes (e.g. [39]) could be an interesting alternative to explore. The idea of a dynamic, stochastic tracking model should then provide the base for generating rainfall scenarios over a relatively vast area of West Africa in order to identify the main sources of variability in the monsoon phenomenon.

7.6. Land Use and Transport Models Calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Arthur Vidard, Peter Sturm, Elise Arnaud.

Given the complexity of modern urban areas, designing sustainable policies calls for more than sheer expert knowledge. This is especially true of transport or land use policies, because of the strong interplay between the land use and the transportation systems. Land use and transport integrated (LUTI) modelling offers invaluable analysis tools for planners working on transportation and urban projects. Yet, very few local authorities in charge of planning make use of these strategic models. The explanation lies first in the difficulty to calibrate these models, second in the lack of confidence in their results, which itself stems from the absence of any well-defined validation procedure. Our expertise in such matters will probably be valuable for improving the reliability of these models. To that purpose we participated to the building up of the ANR project CITiES led by the STEEP EPI. This project started early 2013 and two PhD about sensitivity analysis and calibration were launched late 2013. This work led to conference papers [24], [23] and a two published journal paper [3], [6]

AMIB Project-Team

6. New Results

6.1. RNA Design

In collaboration with J. Hales, J. Manuch and L. Stacho (Simon Fraser University/Univ. British Columbia, Canada), we have investigated the combinatorial RNA design problem, a minimal instance of the RNA design problem which aims at finding a sequence that admits a given target as its unique base pair maximizing structure. We obtained provide complete characterizations for the structures that can be designed using restricted alphabets. We provided a complete characterization of designable structures without unpaired bases. When unpaired bases are allowed, we provides partial characterizations for classes of designable/undesignable structures, and showed that the class of designable structures is closed under the stutter operation. Membership of a given structure to any of the classes can be tested in linear time and, for positive instances, a solution could be found in linear time. Finally, we considered a structure-approximating version of the problem that allows to extend helices and, assuming that the input structure avoids two motifs, we provided a linear-time algorithm that produces a designable structure with at most twice more base pairs than the input structure, as illustrated by Fig. 3 .

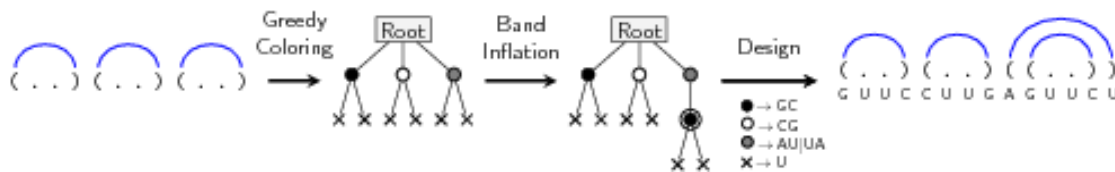


Figure 3. Principle of our structure-approximating version of RNA design: Starting from a potentially undesignable structure, a greedy coloring can be performed and corrected such that the final structure is provably designable in linear time.

Theses results were presented at the CPM 2015 conference in Italy [17], and open new avenues of research, both towards practical, tractable versions of design, and constitute a first step towards long-awaited theoretical foundations for the problem.

6.2. Combinatorics of motifs and algorithms

We developed an $O(n)$ -time and $O(n)$ -space algorithm to compute minimal absent words. Their computation is used in sequence comparison [32] or to detect biologically significant events. For instance, in [52], it was shown that there exist three minimal words in *Ebola* virus genomes which are absent from human genome. The identification of such species-specific sequences may prove to be useful for the development of both diagnosis and therapeutics. In our new contribution [21] we provided an implementation that can be executed in parallel. Experimental results show that excluding the indexing data structure construction time, it achieves near-optimal speed-ups. The computation on the human genome is accelerated by a factor of 10 when using 16 processors, but it consumes a huge amount of RAM. Thus we are currently working on an external memory implementation, that will provide a trade-off between space and time consumption.

Combinatorial tools have been developed to predict the length of repetitions in a random sequence. This allows to distinguish biologically significant repetitions or tune some parameters in assembly or re-sequencing algorithms. For instance, unique mappability is strongly related to the length of the repetitions. A *trie profile* was defined in [45] to address this issue for binary alphabets, by the means of analytic combinatorics. General alphabets, where no closed formula exist, were addressed in [24]. An alternative, and simpler, approach is derived, that exhibits a Large deviation Principle and makes use of Lagrange multipliers. Different domains and transition phases are exhibited. It is expected that this approach extends to a Markov model and to approximate repetitions.

6.3. Structural variants

D. Iakovishina defended in 2015 a PhD thesis co-advised by M. Régnier and V. Boeva (Curie Institute). She proposed a new computational method to detect structural variants using whole genome sequencing data. It combines two techniques that are based either on the detection of paired-end mapping abnormalities or on the detection of the depth of coverage. SV-BAY relies on a probabilistic Bayesian approach and includes a modelization of possible sequencing errors, read mappability profile along the genome and changes in the GC-content. Keeping only somatic SVs is an additional option when matched normal control data are provided. SV-BAY compares favorably with existing tools on simulated and experimental data sets [12] Software SV-BAY is freely available <https://github.com/InstitutCurie/SV-Bay>.

As a side product, a novel exhaustive catalogue of SV types -to date the most comprehensive SV classification- was built. On the grounds of previous publications and experimental data, seven new SV types, ignored by the existing SV calling algorithms, were exhibited.

Structural variations can also be observed and analyzed at larger time scales, and computational methods can be used to predict the structure of ancestral genomes. Within two collaborations with C. Chauve, A. Rajaraman (Simon Fraser University, Canada) and J. Zanetti (SFU, Canada & UniCAMP, Brazil), we revisited the problem of predicting a parsimonious set of adjacencies between ancestral genes, i.e. the most likely structure of an ancestral genome. More specifically, we modified the dynamic programming scheme underlying the DeCo algorithm [28] to compute indicators of robustness for predicting adjacencies. Our reimplementation, which relies on interesting meta-programming strategies, is available at <https://github.com/yannponty/DeClone>.

In a first study, we postulated a Boltzmann-Gibbs distribution over the set of evolutionary scenarios [9]. Our initial experiments relied on Boltzmann sampling to estimate the probabilities of ancestral adjacencies, but our extended version describes an exact polynomial-time computation of such probabilities, through an adaptation of the inside-outside algorithm. We interpreted such probabilities as supports for predicted adjacencies, and found that discarding adjacencies associated with low supports provided a good strategy for resolving synthetic conflicts.

However, the costs associated with the main operations (gaining/breaking adjacencies) in the underlying evolutionary models must be set beforehand in a somewhat arbitrary fashion. This has led us to investigate the influence of those costs on the characteristics of parsimonious predictions, i.e. the robustness of predictions with respect to perturbations of the scoring scheme [18]. To that purpose, we have performed an exact parametric analysis of the DeCo dynamic programming scheme (see Fig. 4 for details). This analysis revealed a quasi-independence, for a large subset of gene trees, of predicted adjacencies to the actual numerical values involved in the scoring scheme.

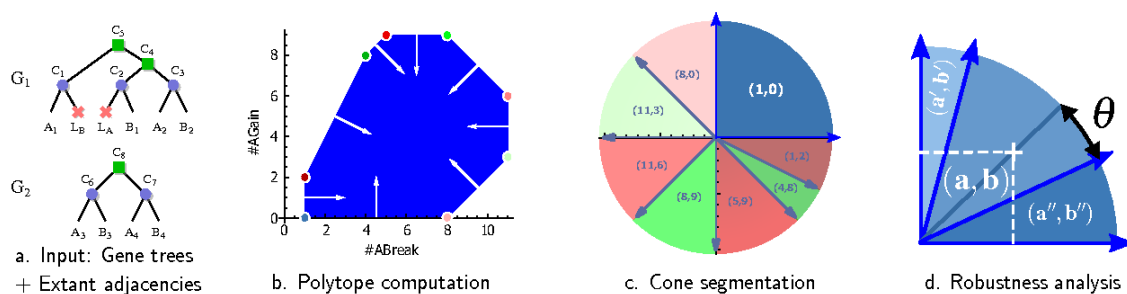


Figure 4. Main steps involved in the parametric prediction of ancestral adjacencies. Starting from two reconciled gene trees and a list of contemporary adjacencies (a.), the polytope of admissible Adjacency Gains/Breaks (+Presence/Absence of a given adjacency) is computed (b.) and projected onto a dual space which partitions the space of cost schemes into (infinite) regions leading to equivalent predictions (c.). The angular distance of the reference cost scheme $(1, 1)$ to a region representing an alternative prediction (d.) is used as a measure of robustness for the prediction.

ANGE Project-Team

7. New Results

7.1. Modelling of complex flows

7.1.1. *Non-hydrostatic models*

Participant: Martin Parisot.

A new shallow water type model involving non-hydrostatic effects is derived in [37]. Under the assumption that the horizontal velocity is close to its vertical mean value, the model enables to recover the energy from the Euler system before integration. Link with the non-hydrostatic published in [18] is identified. Compared to the aforementioned models, the new system consists of more equations (6). However, the numerical strategy presented in the paper does not induce extra computational time.

7.1.2. *Seismic activities: energy radiated by elastic waves*

Participants: Anne Mangeney, Jacques Sainte-Marie.

Estimating the energy loss in elastic waves during an impact is an important problem in seismology and in industry. Three complementary methods to estimate the elastic energy radiated by bead impacts on thin plates and thick blocks from the generated vibration are proposed in [30]. The first two methods are based on the direct wave front and are shown to be equivalent. The third method makes use of the diffuse regime. These methods are shown to be relevant to establish the energy budget of an impact. The radiated elastic energy estimated with the presented methods is quantitatively validated by Hertz's model of elastic impact.

7.1.3. *Layer-averaged Euler and Navier-Stokes systems*

Participants: Marie-Odile Bristeau, Bernard Di Martino, Cindy Guichard, Jacques Sainte-Marie.

In [25] we propose a strategy to approximate incompressible free surface Euler and Navier-Stokes models. The main advantage of the proposed models is that the water depth is a dynamical variable of the system and hence the model is formulated over a fixed domain.

The proposed strategy extends previous works approximating the Euler and Navier-Stokes systems using a multilayer description. Here, the needed closure relations are obtained using an energy-based optimality criterion instead of an asymptotic expansion. Moreover, the layer-averaged description is successfully applied to the Navier-Stokes system with a general form of the Cauchy stress tensor.

7.2. Applications to marine energies

7.2.1. *Partially free surface flow*

Participants: Martin Parisot, Fabien Wahl.

In view of taking into account interactions with buoys, a new formulation of the shallow water model is derived with a constraint corresponding to a static roof. A relaxation approach is considered to adapt the standard numerical schemes. A particular attention is paid to the energy law whether it be for the original model with constraint or the relaxed version.

7.2.2. *Swell energy*

Participants: Sebastian Reyes-Riffo, Julien Salomon.

The internship consisted in designing an optimisation algorithm to determine advantageous topographies in view of producing energy from swell. This approach corresponds to the coupling between a shallow water type model with iterative updates of the topography. Stability of the numerical scheme is a critical point and requires the tuning of parameters.

7.3. Analysis of models in Fluid Mechanics

7.3.1. Weak solutions of multilayer models

Participants: Bernard Di Martino, Ethem Nayir, Yohan Penel.

Proving the existence of global weak solutions is a difficult problem for Navier-Stokes type equations, particularly in case of a degenerate viscosity (viscosity term can vanish if density or thickness goes to zero). In some recent works, Vasseur and Yu [46], have proved this existence for 2D shallow water equations. For the multilayer model, a collaboration with Boris Haspot (Univ. Paris-Dauphine) lead to stability results for the system with a focus on the difficulty to construct a sequence of approximate solutions that conserve all a priori estimates.

7.3.2. Strong solutions of multilayer models

Participants: Emmanuel Audusse, Ethem Nayir, Yohan Penel.

The existence and uniqueness of strong solutions of the multilayer model proposed in [41] was previously proven in the case of boundary conditions. We extended this result to an unbounded domain for short times, overcoming the issue of integrability often barely evoked in similar investigations. Current works deal with the long time existence by a continuation process which requires a particular care of the short time solution at the end of its existence interval.

7.3.3. Hyperbolic problems under constraints

Participant: Nicolas Seguin.

In [21], we study a family of linear hyperbolic systems whose solution must satisfy a constraint (e.g. a simplified model of river flows taking risk of flooding into account). We analyse relaxed models based on a penalisation. This theoretical approach could be used to derive numerical methods.

7.3.4. Entropy-satisfying finite volume schemes

Participant: Nicolas Seguin.

In [44], we carry out an analysis of 1st-order entropy-satisfying finite volume schemes for hyperbolic systems. More precisely, we investigate the numerical dissipation on unstructured meshes under relevant stability conditions. This results in a minimal convergence order towards smooth solutions.

7.3.5. Global existence for Green-Naghdi type equations

Participant: Dena Kazerani.

In [31], we consider the Cauchy problem for the Green-Naghdi equations with viscosity, for small initial data. It is well-known that adding a second order diffusion term to a hyperbolic system leads to the existence of global smooth solutions, as soon as the hyperbolic system is symmetrizable and the so-called Kawashima-Shizuta condition is satisfied. In a previous work, we have proved that the Green-Naghdi equations can be written in a symmetric form, using the associated Hamiltonian. This system being dispersive, in the sense that it involves third order derivatives, the symmetric form is based on symmetric differential operators. We use this structure for an appropriate change of variable to prove that adding viscosity effects through a second order term leads to global existence of smooth solutions, for small data. We also deduce that constant solutions are asymptotically stable.

7.4. Numerical methods for free-surface flows

7.4.1. Godunov schemes for the low Froude regime

Participants: Emmanuel Audusse, Do Minh Hieu, Yohan Penel.

We investigated in [29] the behaviour of collocated Godunov type finite volume schemes when applied to the 1d linear wave equation with Coriolis force in collaboration with S. Dellacherie and P. Omnes (CEA). Accuracy for short time and stability were proven for different versions of the classical Godunov schemes, including some schemes already proposed in the literature (Bouchut *et al.*, [42]). Next step will be to include linear advection and then to study the fully non linear shallow water model. Then results will be extended to 2d problems for which geometrical constraints should be taken into account.

7.4.2. Numerical method for non-hydrostatic models

Participants: Nora Aïssiouene, Marie-Odile Bristeau, Edwige Godlewski, Jacques Sainte-Marie.

In [1], a numerical method based on a prediction-correction scheme in one dimension has been developed and compared to experimental data and analytical solutions. The issue is then to extend the method in higher dimensions. We propose a variational framework for the resolution of a non-hydrostatic Saint-Venant type model with bottom topography. This model is a shallow water type approximation of the free surface incompressible Euler system and slightly differs from the Green-Naghdi model. The resolution of the incompressibility constraint leads to an elliptic problem involving the non-hydrostatic part of the pressure. This step uses a variational formulation of a shallow water version of the incompressibility condition. Several numerical experiments are performed to confirm the relevance of our approach. This work is exposed in [18].

7.4.3. Uncertainties with the topography

Participants: Emmanuel Audusse, Nicole Goutal, Philippe Ung.

We propose to study the uncertainty related to the Saint-Venant system. A perturbation is introduced in the bottom topography such that the topography deviation is characterized by two parameters: its amplitude and its smoothness. In particular, we extend the work previously done with periodic boundary conditions and suggest a treatment of the physical ones. In doing so, we are interested in the influence of the topography deviation on the hydraulic quantities, and in particular, we numerically exhibit a relationship between the spatial correlations of the topography and the water height. Furthermore, we complete the study by a comparison of the outputs between the two flow regimes – fluvial and torrential.

7.4.4. Coupled Stokes-Exner model

Participant: Nora Aïssiouene.

In the framework of the 2015 CEMRACS session (Coupling Multi-Physics Models involving Fluids), we explored an approach to model the sediment transport. In [17], we consider a coupling between the Exner equation and the Stokes system to model sediments in geophysical flow phenomena. We focus on a model without free surface and used some numerical tests to evaluate the relevance of the method. The fluid structure interaction theory and methods have been applied on the coupled system and the objective is to test the proposed method which can be extended to a free surface model. The library Feel++ and the high computing performance embedded have been used to test the solution method. Therefore, the goal of this project is to understand the impact of the sediment transport on the flow using Navier-Stokes with a free surface system coupled with the standard Exner equation. This work has been done in collaboration with Tarik Amtout, Matthieu Brachet, Emmanuel Frenod, Romain Hild, Christophe Prud'homme, Antoine Rousseau and Stéphanie Salmon.

7.5. Software developments and assessments

7.5.1. Improvements in the FRESHKISS3D code

Participants: Marie-Odile Bristeau, David Froger, Raouf Hamouda, Jacques Sainte-Marie.

Several tasks have been achieved in the FRESHKISS3D software:

- The parallelisation of FRESHKISS3D with MPI is achieved for the Eulerian description and the explicit time scheme.
- The paddle wheel vertical effect is now taken into account.
- Vertical and time dependent flow rates can be customised.
- Unit tests have been improved and functional tests have been added.
- Software dependencies are packaged in SED-Paris repository.
- Online documentation is being written.
- A prototype of the software implemented in Cython is under discussion.
- Code executing time's loop is being refactored into multiple classes.
- Various improvements (build system, continuous integrations, coding rules) have been provided.

ARAMIS Project-Team

7. New Results

7.1. Learning spatiotemporal trajectories from manifold-valued longitudinal data

Participants: Jean-Baptiste Schiratti [Correspondant], Stéphanie Allasonniere, Olivier Colliot, Stanley Durrleman.

We propose a Bayesian mixed-effects model to learn typical scenarios of changes from longitudinal manifold-valued data, namely repeated measurements of the same objects or individuals at several points in time. The model allows the estimation of a group-average trajectory in the space of measurements. Random variations of this trajectory result from spatiotemporal transformations, which allow changes in the direction of the trajectory and in the pace at which trajectories are followed. The use of the tools of Riemannian geometry allows to derive a generic algorithm for any kind of data with smooth constraints, which lie therefore on a Riemannian manifold. Stochastic approximations of the Expectation-Maximization algorithm is used to estimate the model parameters in this highly non-linear setting.

The method is used to estimate a data-driven model of the progressive impairments of cognitive functions during the onset of Alzheimer's disease. Experimental results show that the model correctly put into correspondence the age at which each individual was diagnosed with the disease, thus validating the fact that it effectively estimated a normative scenario of disease progression. Random effects provide unique insights into the variations in the ordering and timing of the succession of cognitive impairments across different individuals.

More details in [30] and [31].

7.2. Joint Morphometry of Fiber Tracts and Gray Matter structures using Double Diffeomorphisms

Participants: Pietro Gori [Correspondant], Olivier Colliot, Linda Marrakchi-Kacem, Yulia Worbe, Alexandre Routier, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

This work proposes an atlas construction method to jointly analyse the relative position and shape of fiber tracts and gray matter structures. It is based on a double diffeomorphism which is a cascade of two diffeomorphisms. The first deformation acts only on the white matter keeping fixed the gray matter of the atlas. The resulting white matter, together with the gray matter, are then deformed by the second diffeomorphism which puts into correspondence the homologous anatomical structures across subjects. The first diffeomorphism makes the fiber bundles slide on the fixed gray matter revealing the variability in structural connectivity within the population, namely both the changes in the connected areas and in the geometry of the pathway of the tracts. Fiber bundles are approximated with weighted prototypes using the metric of weighted currents. The algorithm is based on a Bayesian framework which allows the automatic estimation of the covariance matrix of deformation parameters and of the noise variance of each structure. This approach is applied to patients with Tourette syndrome and controls showing a variability in the structural connectivity of the left cortico-putamen circuit.

More details in [26].

7.3. Bayesian Mixed Effect Atlas Estimation with a Diffeomorphic Deformation Model

Participants: Stanley Durrleman [Correspondant], Stéphanie Allasonniere, Estelle Kuhn.

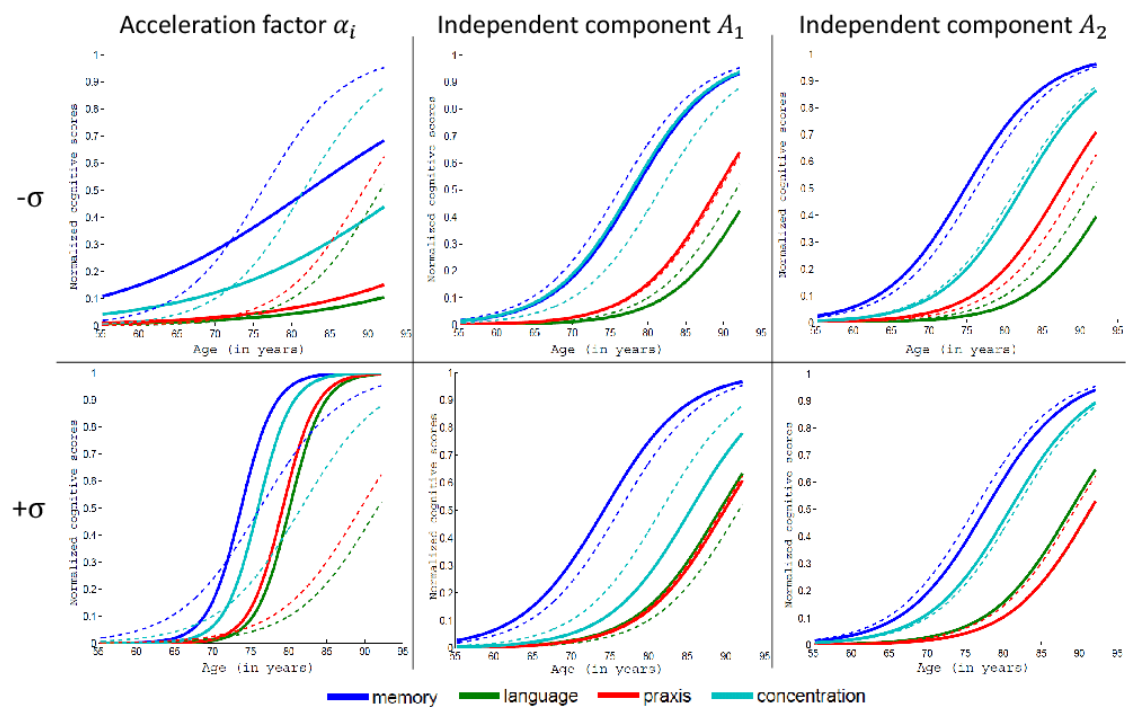


Figure 1. Disease progression model obtained from neuropsychological assessments of 248 patients observed at multiple times (from 3 to 11 times) who converted from Mild Cognitive Impairment stage to Alzheimer's disease during the observation period. Dashed lines represent the average scenario of disease progression (same in all plots). Solid lines represent the variability of this scenario within the observed population in terms of pace of disease progression (left) and relative timing and ordering of the decline of cognitive functions (middle and right).

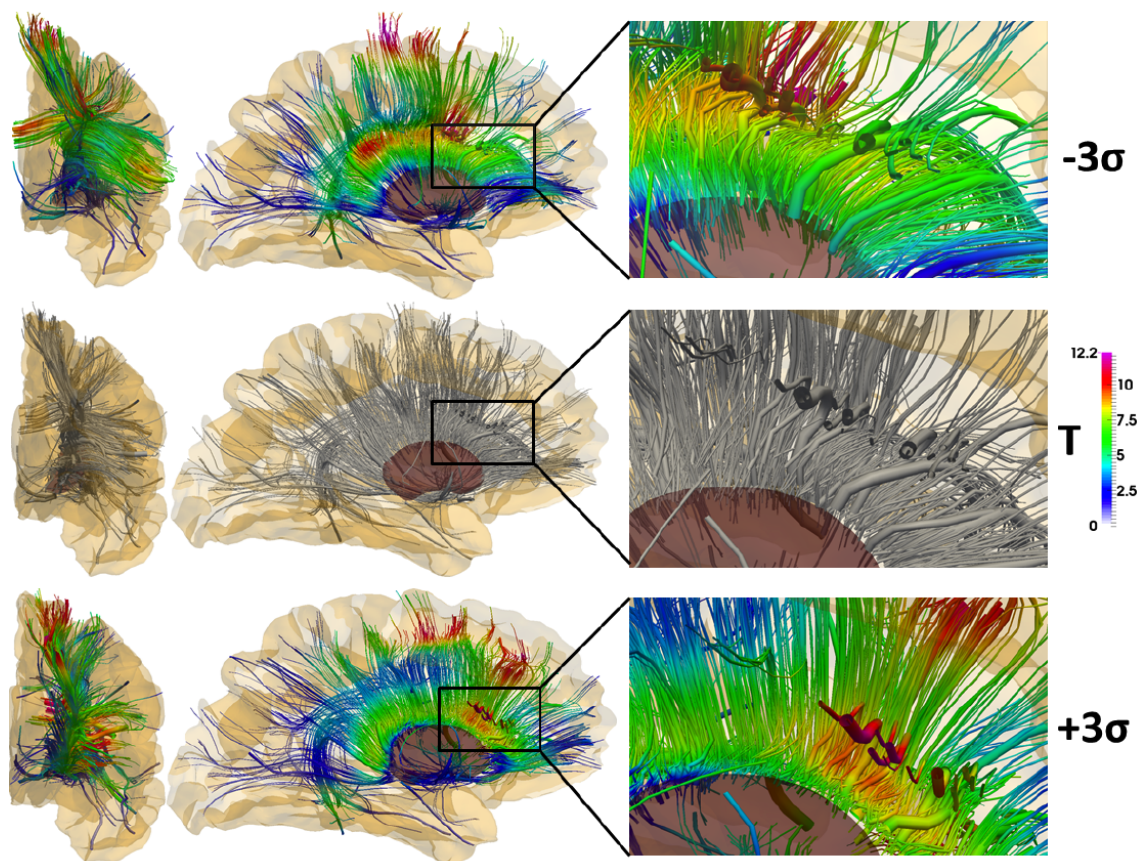


Figure 2. Estimation of a virtual representation of brain structure from anatomical and diffusion images of 3 patients with Gilles de la Tourette syndrome and 2 control subjects. Deformation of the white matter fiber bundle along the first mode of variability is shown while the estimated grey matter frame is kept fixed. Colors refer to the magnitude of displacement during deformation.

In this work, we introduced a diffeomorphic constraint on the deformations considered in the deformable Bayesian Mixed Effect (BME) Template model. Our approach is built on a generic group of diffeomorphisms, which is parameterized by an arbitrary set of control point positions and momentum vectors. This enables us to estimate the optimal positions of control points together with a template image and parameters of the deformation distribution which compose the atlas. We propose to use a stochastic version of the Expectation-Maximization (EM) algorithm where the simulation is performed using the Anisotropic Metropolis Adjusted Langevin Algorithm (AMALA). We propose also an extension of the model including a sparsity constraint to select an optimal number of control points with relevant positions. Experiments are carried out on the USPS database, on mandibles of mice, and on 3D murine dendrite spine images.

More details in [2].

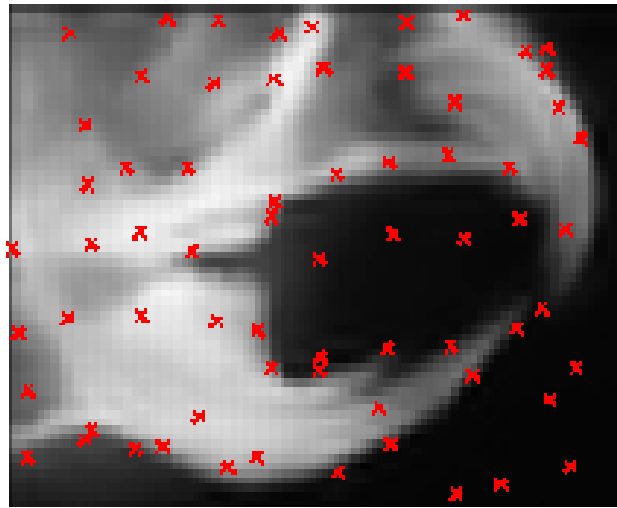


Figure 3. Template image of mouse mandible obtained from 36 X-ray image using 70 control points.

7.4. A sub-Riemannian modular approach for diffeomorphic deformations

Participants: Barbara Gris [Correspondant], Stanley Durrleman, Alain Trouvé.

We develop a generic framework to build large deformations from a combination of base modules. These modules constitute a dynamical dictionary to describe transformations. The method, built on a coherent sub-Riemannian framework, defines a metric on modular deformations and characterises optimal deformations as geodesics for this metric. We present a generic way to build local affine transformations as deformation modules, and display examples.

More details in [27].

7.5. Results of a multicenter randomized placebo-controlled clinical trial in prodromal Alzheimer's disease

Participants: Bruno Dubois, Marie Chupin, Harald Hampel, Simone Lista, Enrica Cavedo, Bernard Croisille, Guy Louis Tisserand, Jacques Touchon, Alain Bonafé, Pierre-Jean Ousset, Amir Ait Ameer, Olivier Rouaud,

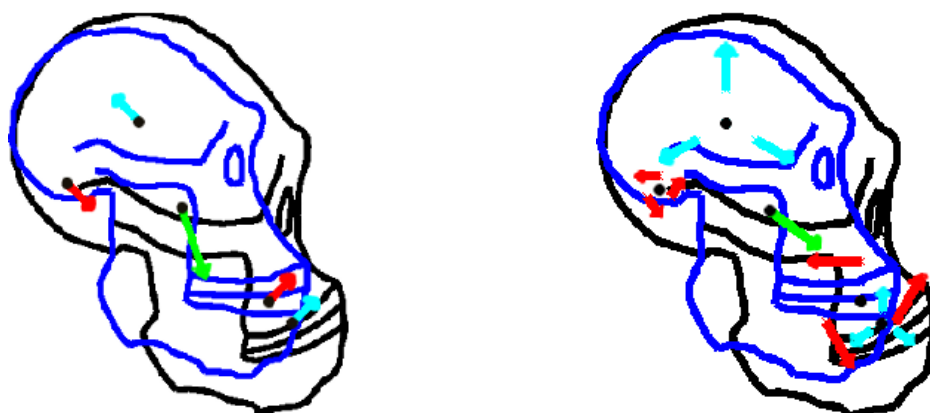


Figure 4. Initial position of deformation modules and their control parameters (left) leads to the construction of local scaling (cyan), rotation (red) and translation (green) (right), which combine together to deform the blue shape into the black one.

Frédéric Ricolfi, Alain Viguetto, Florence Pasquier, Christine Delmaire, Mathieu Ceccaldi, Nadine Girard, Carole Dufouil, Stéphane Lehericy, Isabelle Tonelli, Françoise Duveau, Olivier Colliot, Line Garnero, Marie Sarazin, Didier Dormont [Correspondant].

Our team coordinated neuroimage acquisition and analysis of a multicenter randomized placebo-controlled clinical trial aiming to assess the efficacy of donepezil in prodromal Alzheimer's disease. Subjects underwent two brain magnetic resonance imaging scans (baseline and final visit). The primary efficacy outcome was the annualized percentage change (APC) of total hippocampal volume (left + right) measured by the software (see Section SACHA 6.3) developed by our team. Two-hundred and sixteen only subjects were randomized across 28 French expert clinical sites. In the per protocol population (placebo = 92 and donepezil = 82), the donepezil group exhibited a significant reduced rate of hippocampal atrophy (APC= -1.89%) compared with the placebo group (APC= -3.47%), $P < .001$. There was no significant difference in neuropsychological performance between treatment groups. A 45% reduction of rate of hippocampal atrophy was observed in prodromal AD following 1 year of treatment with donepezil compared with placebo.

This new approach opens interesting perspectives for the evaluation of treatments in neurodegenerative diseases.

More details in [12].

7.6. Sulcal morphology as a new imaging marker for the diagnosis of early onset Alzheimer's disease

Participants: Lorraine Hamelin, Bruno Dubois, Marie Chupin, Olivier Colliot [Correspondant], Marie Sarazin.

We investigated the utility of sulcal width measures in the diagnosis of Alzheimer's disease (AD). Sixty-six biologically confirmed AD patients (positive amyloid positron emission tomography [PET] and/or AD cerebrospinal fluid profile) were contrasted to 35 controls with negative amyloid PET. Patients were classified into prodromal or dementia stages as well as into late onset (LOAD, $n = 31$) or early onset (EOAD, $n = 35$) subgroups according to their age of onset. An automated method was used to calculate sulcal widths and hippocampal volumes (HV). In EOAD, the greatest ability to differentiate patients from age-matched controls, regardless of severity, was displayed by sulcal width of the temporoparietal cortex. In this region, diagnosis

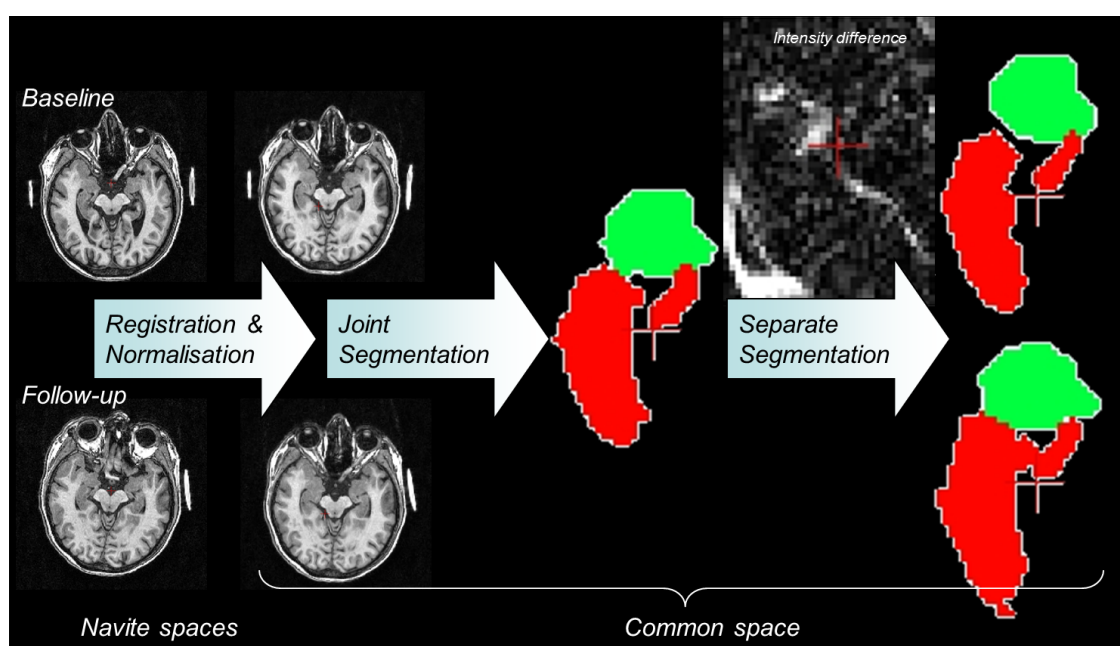


Figure 5. Hippocampus longitudinal segmentation method illustrating preliminary registration of the baseline and final visit magnetic resonance imaging (MRI) scans in a common space followed by normalization of the intensities of both scans. The baseline and final visit MRI scans were then segmented jointly. The resulting segmentation was then used as an initialization of separate segmentations while keeping the two segmentations consistent between the two time-points.

accuracy was better than the HV, especially at prodromal stage. In LOAD, HV provided the best discrimination power from age-matched controls. In conclusion, sulcal width measures are better markers than the HV for identifying prodromal AD in patients aged <65 years. In contrast, in older patients, the risk of over-diagnosis from using only sulcal enlargement is important.

More details in [14].

7.7. Imaging Markers of the Presymptomatic GRN Disease

Participants: Paola Caroppo, Stanley Durrleman, Alexandre Routier, Olivier Colliot [Correspondant], Alexis Brice, Isabelle Le Ber.

The preclinical stage of frontotemporal lobar degeneration (FTLD) is not well characterized. We conducted a brain metabolism (FDG-PET) and structural (cortical thickness) study to detect early changes in asymptomatic GRN mutation carriers (aGRN+) that were evaluated longitudinally over a 20-month period. At baseline, a left lateral temporal lobe hypometabolism was present in aGRN+ without any structural changes. Importantly, this is the first longitudinal study and, across time, the metabolism more rapidly decreased in aGRN+ in lateral temporal and frontal regions. The main structural change observed in the longitudinal study was a reduction of cortical thickness in the left lateral temporal lobe in carriers (Figure 6). A limit of this study is the relatively small sample (n=16); nevertheless, it provides important results. First, it evidences that the pathological processes develop a long time before clinical onset, and that early neuroimaging changes might be detected approximately 20 years before the clinical onset of disease. Second, it suggests that metabolic changes are detectable before structural modifications and cognitive deficits. Third, both the baseline and longitudinal studies provide converging results implicating lateral temporal lobe as early involved in GRN disease. Finally, our study demonstrates that structural and metabolic changes could represent possible biomarkers to monitor the progression of disease in the presymptomatic stage toward clinical onset.

More details in [6].

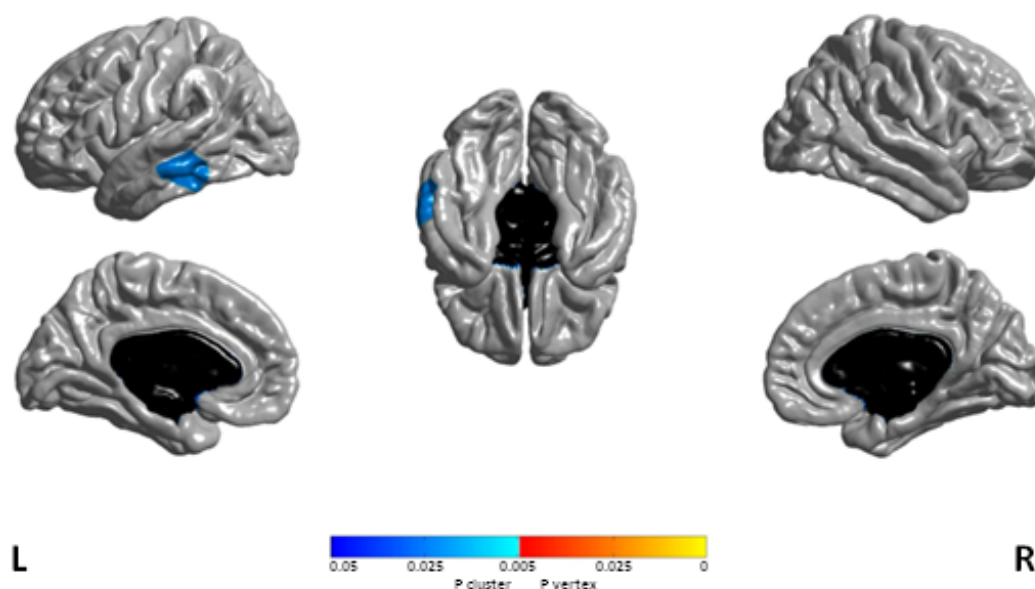


Figure 6. Cluster with significant cortical thickness changes in aGRN+ between the two time-points ($p < 0.05$ corrected). L, left; R, right.

7.8. Incomplete Hippocampal Inversions in healthy subjects: a comprehensive study of over 2000 participants

Participants: Claire Cury [Correspondant], Joan Glaunès, Dominique Hasboun, Fanny Cohen, Jorge Samper-González, Roberto Toro, Vincent Frouin, Gunter Schumann, Olivier Colliot.

The incomplete-hippocampal-inversion (IHI), also known as malrotation, is an atypical anatomical pattern of the hippocampus, which has been reported in healthy subjects in different studies. However, extensive characterization of IHI in a large sample has not yet been performed. Furthermore, it is unclear whether IHI are restricted to the medial-temporal lobe or are associated with more extensive anatomical changes. Here, we studied the characteristics of IHI in a community-based sample of 2008 subjects of the IMAGEN database and their association with extra-hippocampal anatomical variations. The presence of IHI was assessed on T1-weighted anatomical magnetic resonance imaging (MRI) using visual criteria. We assessed the association of IHI with other anatomical changes throughout the brain using automatic morphometry of cortical sulci. We found that IHI were much more frequent in the left hippocampus (left: 17%, right: 6%, χ^2 -test, $p < 10^{-28}$). Compared to subjects without IHI, subjects with IHI displayed morphological changes in several sulci located mainly in the limbic lobe. Our results demonstrate that IHI are a common left-sided phenomenon in normal subjects and that they are associated with morphological changes outside the medial temporal lobe.

More details in [9].

7.9. Analysis of anatomical variability using diffeomorphic iterative centroid in patients with Alzheimer's disease

Participants: Claire Cury [Correspondant], Joan Glaunès, Marie Chupin, Olivier Colliot.

We proposed a new approach for template-based analysis of anatomical variability in populations, in the framework of Large Deformation Diffeomorphic Metric Mappings and mathematical currents. We propose a fast approach in which the template is computed using an diffeomorphic iterative centroid method. Statistical analysis is then performed on the initial momenta that define the deformations between the centroid and each individual subject. We applied the approach to study the variability of the hippocampus in 134 patients with Alzheimer's disease (AD) and 160 elderly control subjects. We show that this approach can describe the main modes of variability of the two populations and can predict the performance to a memory test in AD patients.

More details in [8].

7.10. Innovation-based sparse estimation of functional connectivity from multivariate autoregressive models

Participants: Fabrizio de Vico Fallani [Correspondant], Stéphanie Allasonniere [Correspondant], Francois Deloche.

One of the main limitations of functional connectivity estimators of brain networks is that they can suffer from statistical reliability when the number of areas is large and the available time series are short. To estimate directed functional connectivity with multivariate autoregressive (MVAR) model on sparse connectivity assumption, we propose a modified Group Lasso procedure with an adapted penalty. Our procedure includes the innovation estimates as explaining variables. This approach is inspired by two criteria that are used to interpret the coefficients of the MVAR model, the Directed Transfer Function (DTF) and the Partial Directed Coherence (PDC). A causality measure can be deduced from the output coefficients which can be understood as a synthesis of PDC and DTF. We demonstrate the potential of our method and compare our results with the standard Group Lasso on simulated data.

More details in [25]

7.11. Lucid Dreaming in Narcolepsy

Participants: Pauline Daudet, Mario Chavez [Correspondant], Smaranda Leu-Semenescu, Jean-Louis Golmard, Isabelle Arnulf.

Lucid dreaming is the experience of being aware of dreaming while asleep and continuing to dream. Lucid dreams generally arise in REM sleep. Compared to non-lucid REM sleep, lucid REM sleep is associated with local frontal lobe EEG changes in the 40 Hz band, increased brain coherence, and increased activity on functional MRI in the bilateral precuneus, cuneus, parietal lobules, and prefrontal and occipito-temporal cortices, which may correspond to restored reflective consciousness. We decided to study the frequency and determinants of lucid dreaming in narcolepsy and to challenge patients' alleged ability to achieve lucid dreaming using sleep monitoring during nighttime and daytime sleep. Compared to 53 healthy controls, the 53 narcolepsy patients reported more frequent dream recall, nightmares and recurrent dreams. The frequency of cataplexy, hallucinations, sleep paralysis, dyssomnia, positivity, and the severity of sleepiness were similar in narcolepsy with and without lucid dreaming. The delta power in the electrode average, in delta, theta, and alpha powers in C4, and coherences between frontal electrodes were lower in lucid than non-lucid REM sleep in spectral EEG analysis. The duration of REM sleep was longer, the REM sleep onset latency tended to be shorter, and the percentage of atonia tended to be higher in lucid vs. non-lucid REM sleep; the arousal index and REM density and amplitude were unchanged. Our results suggest that narcoleptics have a high propensity for lucid dreaming without differing in REM sleep characteristics from people without narcolepsy. This also suggests that narcolepsy patients may provide useful information in future studies on the nature of lucid dreaming.

More details in [11]

7.12. An Algebraic Topological Method for Multimodal Brain Networks Comparisons

Participants: Tiago Simas, Mario Chavez [Correspondant], Pablo Rodriguez, Albert Diaz-Guilera.

Understanding brain connectivity is one of the most important issues in neuroscience. Nonetheless, connectivity data can reflect either functional relationships of brain activities or anatomical connections between brain areas. Although both representations should be related, this relationship is not straightforward. We have devised a powerful method that allows different operations between networks that share the same set of nodes, by embedding them in a common metric space, enforcing transitivity to the graph topology. Here, we apply this method to construct an aggregated network from a set of functional graphs, each one from a different subject. Once this aggregated functional network is constructed, we use again our method to compare it with the structural connectivity to identify particular brain regions that differ in both modalities (anatomical and functional). Remarkably, these brain regions include functional areas that form part of the classical resting state networks. We conclude that our method -based on the comparison of the aggregated functional network- reveals some emerging features that could not be observed when the comparison is performed with the classical averaged functional network.

More details in [23]

7.13. Steady state visual evoked potentials-based patient interface under breathing constraints

Participants: Xavier Navarro [Correspondant], Sebastien Campion, Fabrizio de Vico Fallani [Correspondant], Pierre Pouget, Thomas Similowski, Mathieu Raux, Mario Chavez.

Steady state visual evoked potentials (SSVEP) have been widely utilized in brain computer interfacing (BCI) in last years. In this paper, we present a study exploring the possibilities of SSVEP to manage the communication between patients suffering respiratory disorders and health care providers. By imposing different breathing constraints, five healthy subjects communicated their breathing sensations (breathing well/breathing bad) using a visual frequency tagging paradigm: two visual stimuli with different flickering frequencies (15 and 20 Hz) were simultaneously presented on a screen. Using electroencephalographic (EEG) signals from only three EEG electrodes, two spectral features were extracted by a spatial filter in a sliding window, then classified by an unsupervised algorithm based on k-medians. Average detection success rates were of 70% during breathing

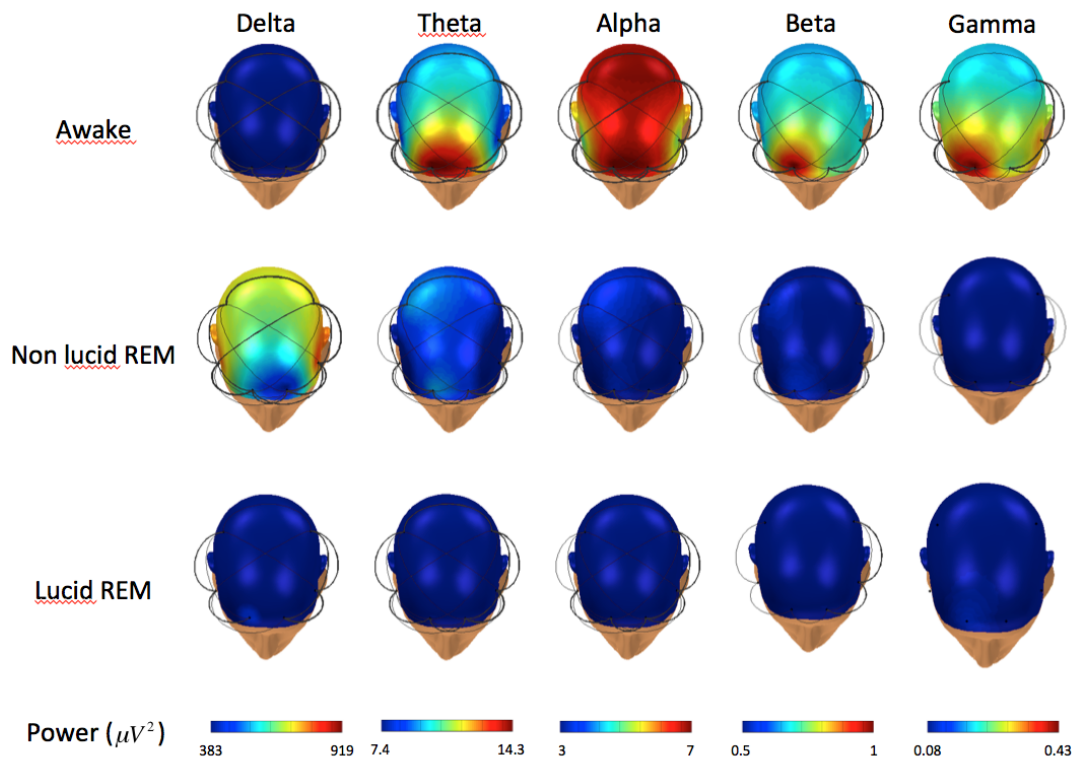


Figure 7. Topographical distribution (obtained by a spherical spline interpolation) of EEG spectral power during wakefulness (top row), non-lucid (middle row) and lucid (bottom row) REM sleep for different frequency bands. Significant couplings between the electrodes are indicated by the black links (the thickness is proportional to the coherence value). Colors from dark blue (lower EEG power) to dark red (higher EEG power) indicated for each EEG band in the Power line (bottom row).

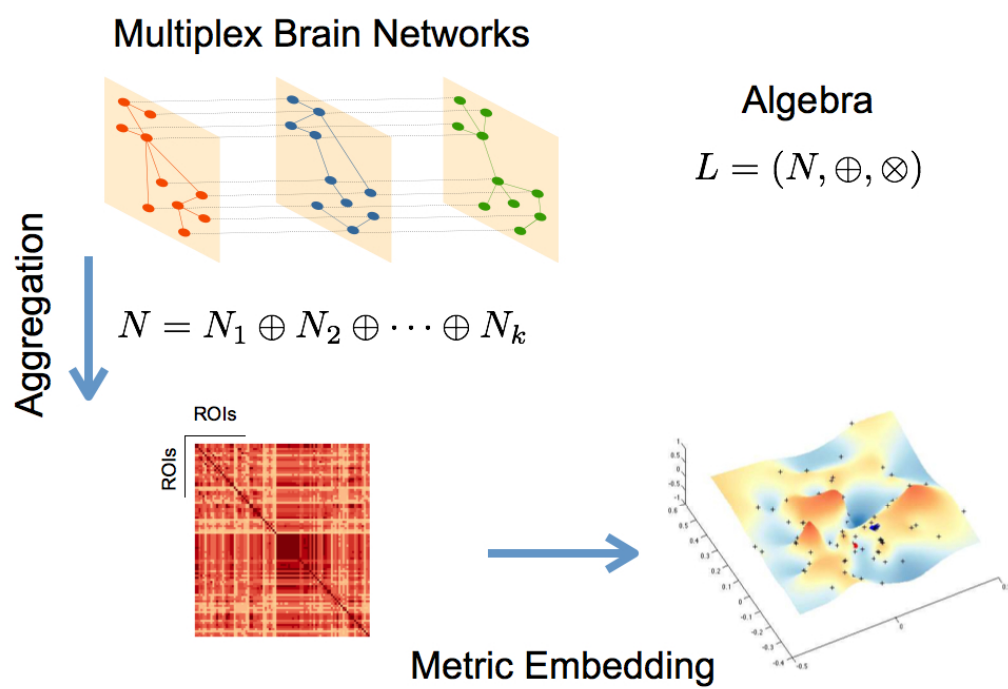


Figure 8. Schematic representation of the main steps for the described networks aggregation and metric embedding (defined here for the algebra L)

discomfort, and of 83% when subjects breathed comfortably. Results suggest that SSVEP-based BCI may be a promising choice to improve patient-caregiver communication in situations of breathing discomfort when verbal communication is difficult.

More details in [29]

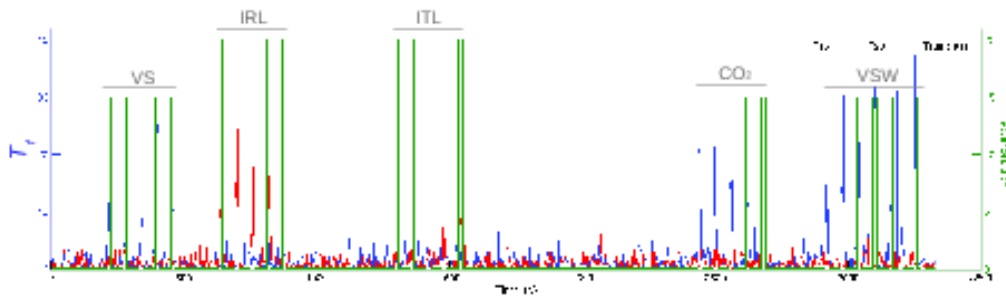


Figure 9. An example of T_f obtained after applying the spatial filters on 15 Hz (blue curve) and 20 Hz (red curve) during the experiment in subject 3. The statistics T_f reflects the signal-to-noise ratio at frequency f with respect to the no-stimulus power.

ASCLEPIOS Project-Team

6. New Results

6.1. Medical Image Analysis

6.1.1. Longitudinal Analysis and Modeling of Brain Development

Participants: Mehdi Hadj-Hamou [Correspondent], Xavier Pennec, Nicholas Ayache, Hervé Lemaître [Inserm U1000], Jean-Luc Martinot [Inserm U1000].

This work is partly funded through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Processing pipeline - brain development - adolescence - longitudinal analysis - non-rigid registration algorithm - extrapolation

1. We proposed and detailed a deformation-based morphometry computational framework, called Longitudinal Log-Demons Framework (LLDF), which estimates the longitudinal brain deformations from image data series, transports them in a common space and performs statistical group-wise analyses (see Fig. 1). This processing pipeline is based on freely available softwares and relies on the LCC log-Demons non-linear diffeomorphic registration algorithm with an additional modulation of the similarity term using a confidence mask to increase robustness with respect to brain boundary intensity artifacts.
2. The LLDF framework is applied to the study of longitudinal trajectories during adolescence, for which little is known. The aim of this project is to provide models of brain development during adolescence based on diffeomorphic registration parametrised by SVFs. Our study focused particularly on the link between sexual dimorphism and the longitudinal evolution of the brain. This work was done in collaboration with J.L. Martinot et H. Lemaître (Inserm U1000).

6.1.2. Inter-Operative Relocalization in Flexible Endoscopy

Participants: Anant Suraj Vemuri [Correspondent], Stéphane Nicolau, Luc Soler, Nicholas Ayache.

This work has been performed in collaboration with IHU Strasbourg and IRCAD, France.

Computer Assisted Intervention, Barrett's Esophagus, Biopsy Relocalization, Electromagnetic tracking

Oesophageal adenocarcinoma arises from Barrett's oesophagus, which is the most serious complication of gastro-oesophageal reflux disease. Strategies for screening involve periodic surveillance and tissue biopsies. A major challenge in such regular examinations is to record and track the disease evolution and relocalization of biopsied sites to provide targeted treatments.

In an earlier paper, we introduced the first approach to inter-operative relocalization using electromagnetic tracking system. In [21], we propose three incremental experiments to our approach. First, we analyse the error bounds of our system on synthetic data with a realistic noise model. Second, we provide a pseudo ground-truth on *in-vivo* pig data using an optical tracking system. Accuracy results obtained were consistent with the synthetic experiments despite uncertainty introduced due to breathing motion, and remain inside acceptable error margins according to medical experts. Finally, a third experiment was designed using data from pigs to simulate a real task of biopsy site relocalization, and evaluated by ten experts. It clearly demonstrated the benefit of our system towards assisted guidance by improving the biopsy site retrieval rate from 47.5% to 94%.

This inter-operative relocalization framework was then further extended in [53] to provide a constrained image based search as shown in Fig. 2 to obtain the best view point match to the live view. Within this context, we investigate the effect of (a) the choice of feature descriptors and colour-space, (b) filtering of uninformative frames and (c) endoscopic modality, for view point localization. Our experiments indicate an improvement in the best view-point retrieval rate to [92%, 87%] from [73%, 76%] (in our previous approach) for Narrow band imaging and white-light endoscopic image modalities.

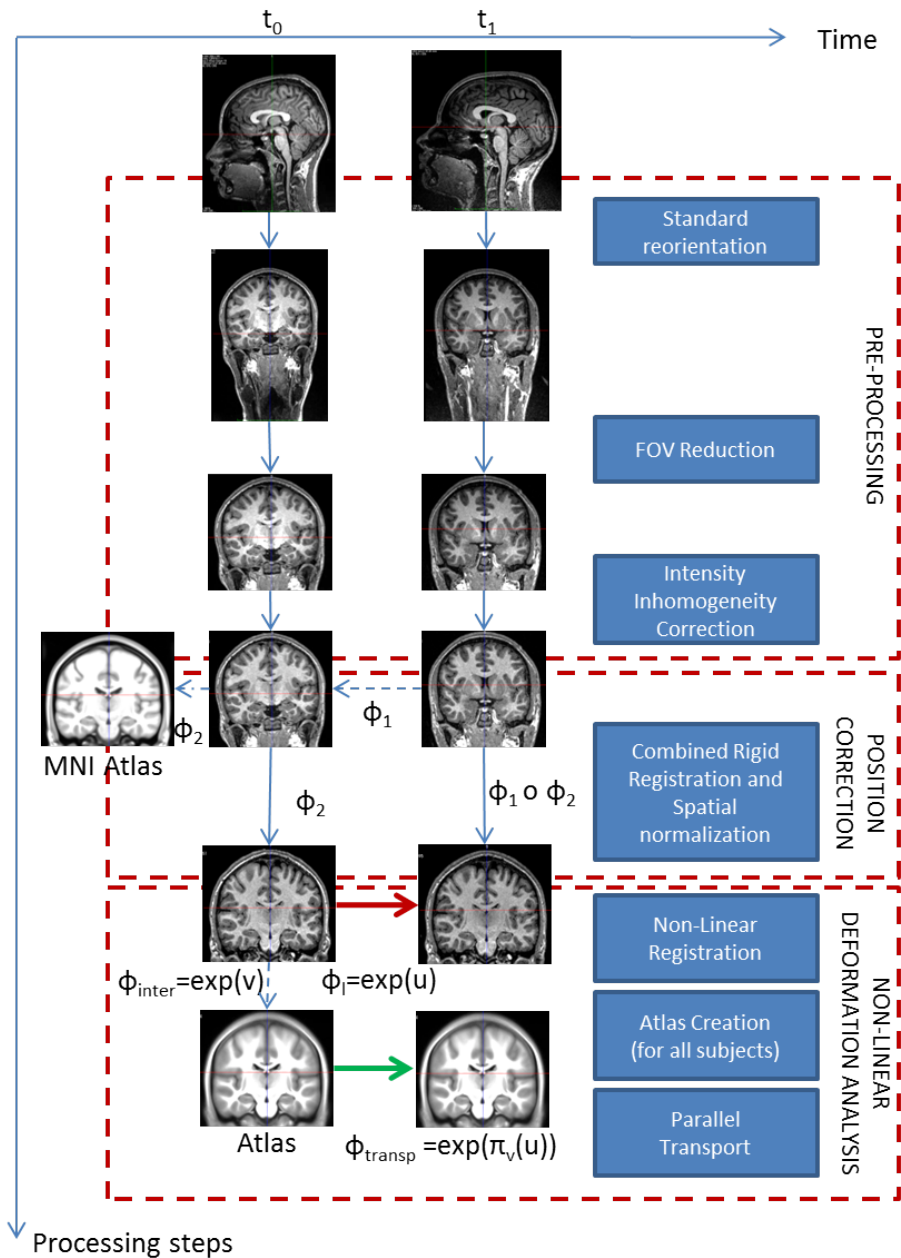


Figure 1. Proposed processing pipeline for longitudinal analysis: the pipeline is composed of three major steps. Starting with raw images, we first pre-process them, then correct the spatial position differences to end up with the longitudinal deformations for each subject in the atlas space.

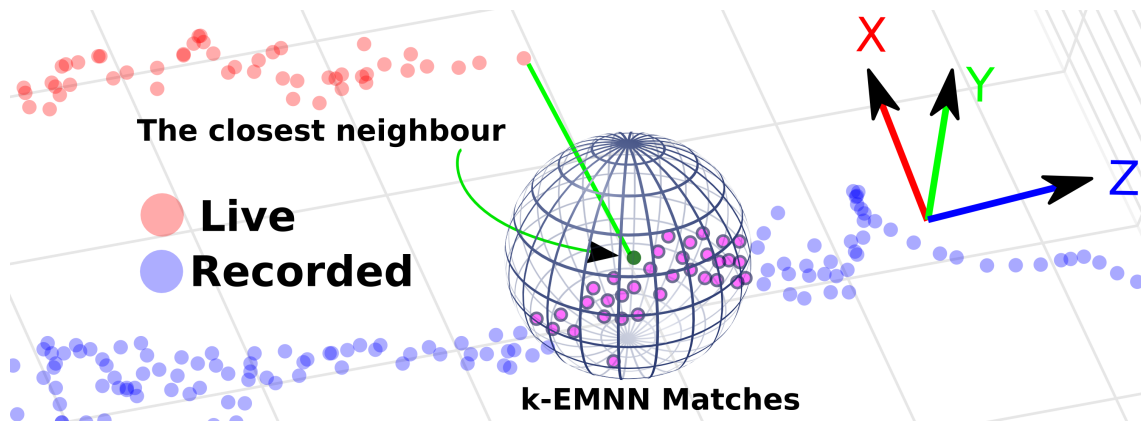


Figure 2. ElectroMagnetic tracker Nearest Neighbour (EMNN) matches. Firstly, the EMNN is obtained using the 3D position match. Then, in a radius r around this match, all the points on the trajectory are considered as the k -EMNN matches. For the images in the k -EMNN matches, scene matching is performed.

6.1.3. Segmentation and anatomic variability of the cochlea and other temporal bone structures from medical images

Participants: Thomas Demarcy [Correspondent], Hervé Delingette, Clair Vandersteen [IUFC, Nice], Dan Gnansia [Oticon Medical], Nicholas Ayache.

This work is supported by the National Association for Research in Technology (ANRT) through the CIFRE Grant 2013-1165 and Oticon Medical (Vallauris). Part of this work is also funded by the European Research Council through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images). This work is done in collaboration with the Department of Ear Nose Throat Surgery (IUFC, Nice) and the Nice University Hospital (CHU).

image segmentation ; surgery planning ; shape modelling ; anatomic variability ; cochlear implant ; temporal bone

- We designed a parametric shape model of the intracochlear anatomy with anatomical prior learned from temporal bones high-resolution images, see Fig. 3 .
- We evaluated the cochleostomy location regarding two surgical approaches (endaural compared to conventional posterior tympanotomy) [20].

6.1.4. Structured sparse Bayesian modelling for non-rigid registration and cardiac motion tracking

Participants: Loic Le Folgoc [Correspondent], Hervé Delingette, Antonio Criminisi, Nicholas Ayache.

This work has been partly supported by the Inria – Microsoft Research Joint Center and by the European Research Council through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Non-rigid Registration - Structured Sparse Bayesian Learning - Automatic Relevance Determination - Reversible-jump Markov Chain Monte Carlo - Cardiac Motion Tracking - Uncertainty Quantification

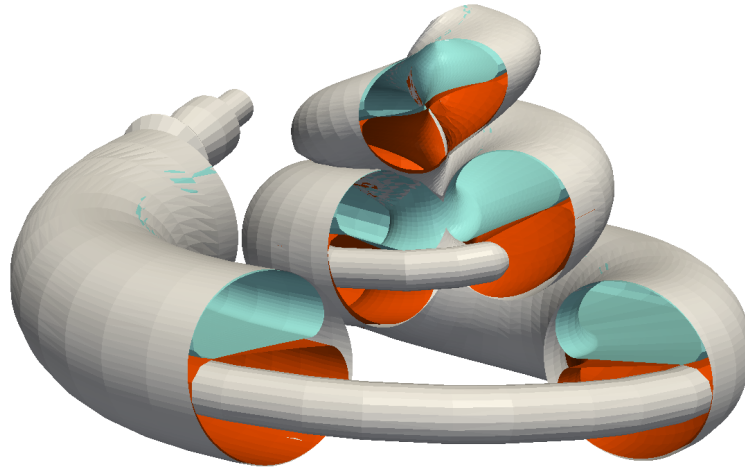


Figure 3. Cochlear implant electrode-array with respect to scala tympani (red) and scala vestibuli (blue).

We developed a generic structured sparse Bayesian model of image registration with three main contributions: an extended image similarity term, the automated tuning of registration parameters and uncertainty quantification. We proposed an approximate inference scheme that is tractable on 4D clinical data. We demonstrated the performance of our approach on cine MR, tagged MR and 3D Ultra Sound cardiac images, and showed state-of-the-art results on benchmark datasets evaluating accuracy of motion and strain.

Moreover, we evaluated the quality of uncertainty estimates returned by the approximate inference scheme. We compare the predictions of the approximate scheme with those of an inference scheme developed on the grounds of reversible jump Markov Chain Monte-Carlo [94](see Fig. 4). We provided more insight into the theoretical properties of the sparse structured Bayesian model and into the empirical behaviour of both inference schemes.

This work is described in the PhD manuscript of Loïc Le Folgoc, defended at Université Nice Sophia Antipolis, 2015 [6].

6.1.5. Image Segmentation and Synthesis of brain tumor MR images

Participants: Nicolas Cordier [correspondent], Hervé Delingette, Nicholas Ayache.

Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Brain, MRI, Glioma, Patch-based Segmentation, Image synthesis

The segmentation of glioblastoma, the most severe case of brain tumors, is a crucial step for diagnostic assessment and therapy planning. In order to perform the manual delineation of the tumor compartments, the clinicians have to concurrently screen multi-channel 3D MRI, which makes the process both time-consuming and subject to inter-expert delineation variability.

We have developed 2 contributions for the analysis of MR brain tumor images:

- A patch-based multi-atlas automatic glioma segmentation algorithm[13]. Unlike prior work on patch-based multi-atlas segmentation, our approach does not assume any prior knowledge about the location of pathological structures (no local search window).

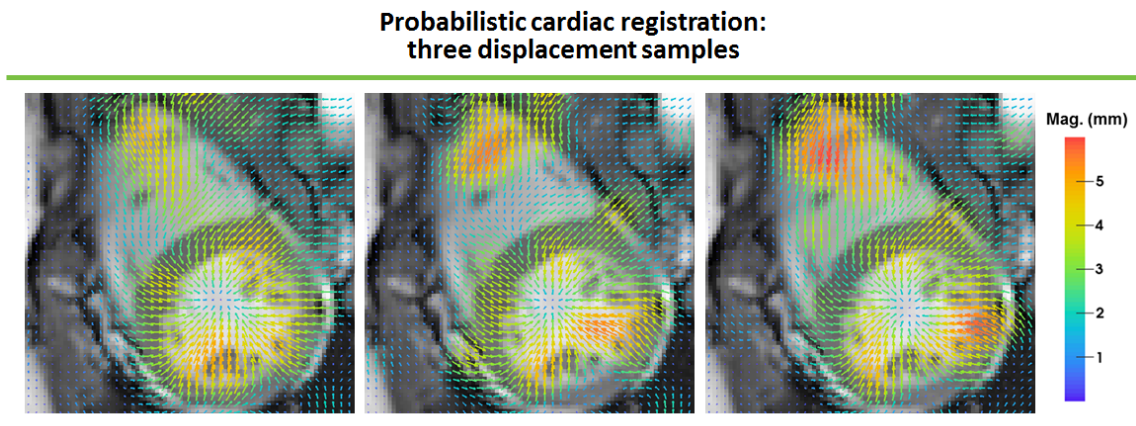


Figure 4. Three displacement fields sampled from the same posterior distribution of coefficients associated with the registration of two cardiac images.

- A patch-based image synthesis algorithm (see Fig.5) [4], which generates multi-sequence MR images of the brain with glioma from a single label image. The synthesis of images may be useful to benchmark segmentation algorithms or to increase the size of annotated medical image databases.

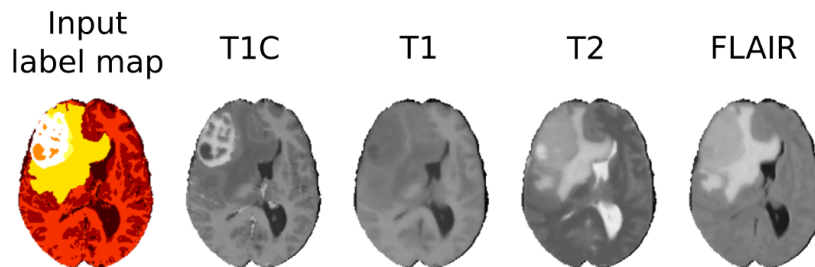


Figure 5. Synthesis of high-grade glioma MR image from a single label map.

6.1.6. Infarct localization from myocardial deformation

Participants: Nicolas Duchateau [Correspondent], Maxime Sermesant.

This work received the partial support from the European Union 7th Framework Programme (VP2HF FP7-2013-611823) and the European Research Council (MedYMA ERC-AdG-2011-291080).

Myocardial infarct, Computer-aided diagnosis, Dimensionality reduction, Biomechanical modeling

- We investigate new methods for predicting the location of myocardial infarcts from local wall deformation [31], which is useful for risk stratification from routine examinations such as 3D echocardiography. In a broader perspective, this project also aims at determining relevant biomarkers to study cardiac function [54], and eventually at combining several of those markers in an efficient manner [59].

- Non-linear dimensionality reduction aims at estimating the Euclidean space of coordinates encoding deformation patterns, and is combined with multi-scale kernel regressions to infer the low-dimensional coordinates and the infarct location of new cases.
- These concepts were tested on 500 synthetic cases with infarcts of random extent, shape, and location, generated from a realistic electromechanical model. Our prediction goes beyond the current diagnosis of infarct either achieved at the global or segmental level, and significantly outperforms the clinically-used thresholding of the deformation patterns.

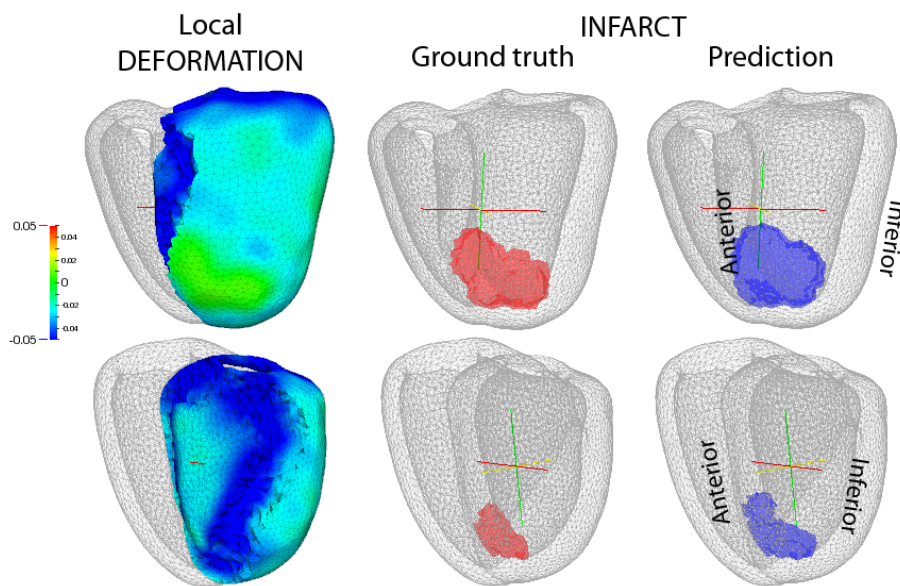


Figure 6. Examples of myocardial deformation patterns, ground truth infarct location, and estimated infarct location.

6.2. Computational Anatomy

6.2.1. Geometric generative model of organ shapes: statistical properties of template shape estimation

Participants: Nina Miolane [Correspondent, Inria - Stanford], Xavier Pennec [Inria], Susan Holmes [Stanford].

This work is conducted jointly with the Department of Statistics of Stanford, in the context of the associated team GeomStats and the FSCIS (France-Stanford Center for Interdisciplinary Studies) fellowship of Nina Miolane.

template, atlas, consistency, estimation theory, Expectation-Maximization algorithm, shapes, quotient space, lie group, sub-Riemannian, in-painting, neuro-geometry, visual cortex, diffusion

This work focuses on the interaction between statistics and geometry, for applications in Medical Imaging. The first part deals with a generative model of (organ) shapes and, more precisely, on the estimation of the mean shape or template. The second part of this work surveys and unveils the mathematical framework needed to extend Neurogeometry, used in 2D Computer Vision, to applications in 3D imaging.

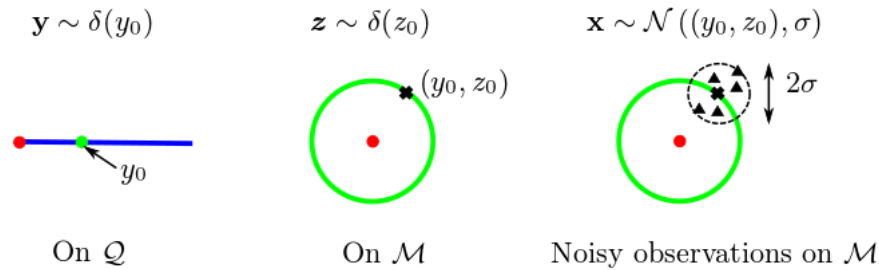


Figure 7. A schematic representation of the generative model of organ shapes.

- In the first part, we define a geometric statistical framework of an organ shapes generative model (see Figure 7). This is done through the differential geometry of quotient spaces.
- Then, we interpret the computation of the mean organ shape (or template) through the max-max algorithm, as an approximate maximum likelihood estimation in this framework.
- Finally, we study the statistical properties of the template computed with this procedure. More precisely, we show that the estimation is inconsistent and that the inconsistency cannot be neglected when the real template is close to the singularity of the quotient space at the scale of the ambient noise on the images [44].
- In the second part, the particularities of a 3D neurogeometry are highlighted with respect to the 2D case. They rely on the fact that 2D neurogeometry is inspired by the primary visual cortex, which codes for our 2D visual field (our retina is 2D). Imagining a 3D visual field or a 3D retina would give rise to a 3D neurogeometry.
- The conceptual framework of a 3D neurogeometry is more subtle, and a new level of mathematical structures arises (see Figure 8). Thus, inpainting (sub-Riemannian diffusion) have to be generalized.
- Applications for in-painting or super-resolution in 3D medical images are described [43].

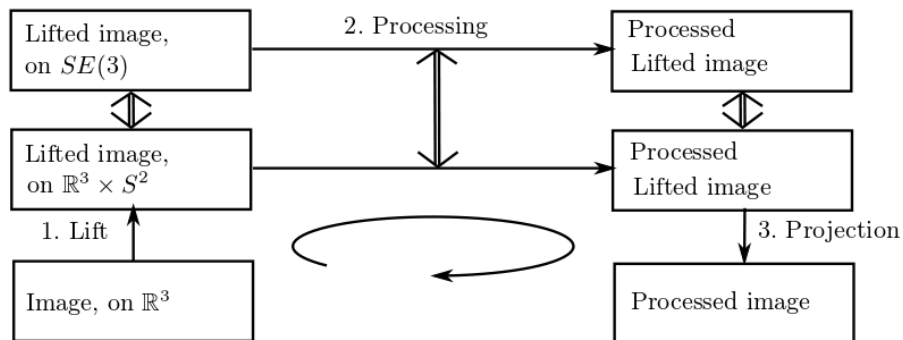


Figure 8. The 3 steps of an image processing pipeline in 3D neurogeometry.

6.2.2. Compact representation of longitudinal deformations

Participants: Raphaël Sivera [Correspondent], Hervé Delingette, Nicholas Ayache.

This work is supported by a PhD fellowship from the University Nice Sophia Antipolis and by the European Research Council through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Longitudinal modeling, Learning in manifolds, Structured sparsity.

The analysis of dynamic or longitudinal series of medical images is important to better understand the observed evolutions of the organs but also to provide robust computer aided diagnosis tools. This analysis can be performed through a reduced representation of geometric transformations capturing the deformation between 2 time points.

In the context of cardiac motion analysis, we proposed a framework to represent arbitrary diffeomorphisms described as Stationary Velocity Fields (SVF) in a low dimensional linear space (see fig. 9).

To this end, we first improved the Inverse Scaling and Squaring (ISS) algorithm from [83] to transform displacement fields into SVFs. Second, through a structured sparse decomposition of these deformations over the cardiac cycle, we provided a preliminary approach for comparing trajectories of cine-MR images between two patients.

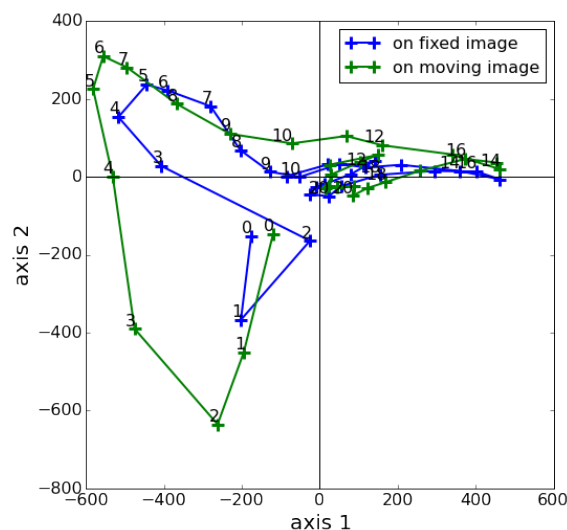


Figure 9. Trajectories of two registered cardiac cycles projected on a 2D space using dimensionality reduction tools.

6.2.3. Statistical analysis of heart motion

Participants: Marc-Michel Rohé [Correspondent], Nicolas Duchateau, Maxime Sermesant, Xavier Pennec.

This work is partly supported by the FP7 European project MD-Paedegree and by the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Statistical analysis, Registration, Reduced order models, Machine learning

This work aims at developing statistical tools to analyze cardiac motion. In particular, we are interested in approximating complex motion models with few parameters or modes that are clinically relevant (reduced models). To this end, we have introduced a polyaffine cardiac motion model that reduces the deformation parameters to a few interpretable parameters, and the most important modes to represent the variability seen in a population are automatically selected. We then performed a group-wise statistical analysis, which relates the model parameters to clinical indices specific to a given pathology. This method was used to classify a population of healthy/infarcted hearts [48] (see Fig. 10), as well as to study cardiac motion of adolescents with cardiomyopathies within the European project "MD-Paedigree".

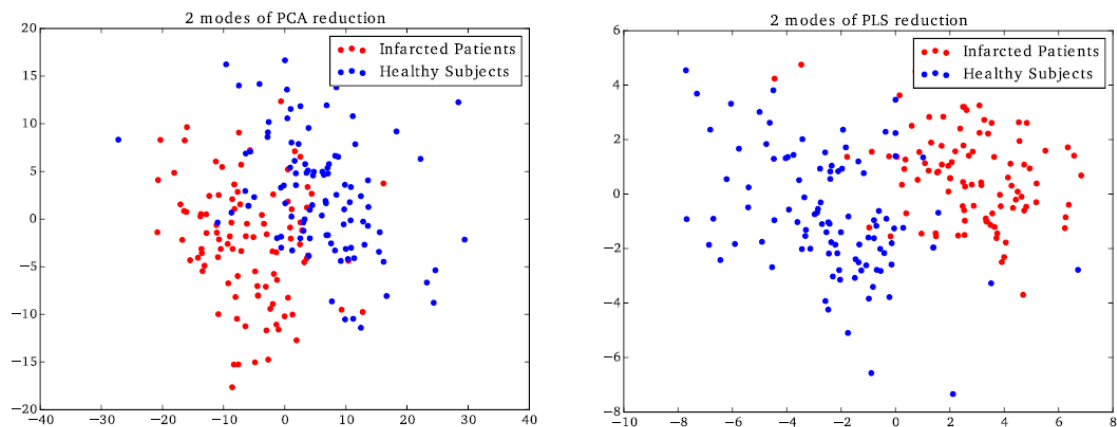


Figure 10. Projection of healthy/infarcted patients cardiac motions on two modes extracted from PCA (left) and PLS (right) methods.

6.2.4. Statistical Learning via Synthesis of Medical Images

Participants: Hervé Lombaert [Correspondent], Héloïse Bleton, Hervé Delingette, Nicholas Ayache, Antonio Criminisi.

This work is partly supported by a grant from Microsoft Research-Inria Joint Centre and by the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

statistical learning, synthesis

Machine learning approaches typically require large training datasets in order to capture as much variability as possible. Application of conventional learning methods on medical images is difficult due to the large variability that exists among patients, pathologies and image acquisitions. The project aims at exploring how realistic image synthesis could be used to improve existing machine learning methods.

We tackled the problem of better exploiting existing training sets, via a smart modeling of the image space, and applying conventional random forests using guided bagging [99]. Synthesis of complex data, such as cardiac diffusion images (DTI), was also done, with a refined version of [98].

Then, we tackled the problem of exploiting *Geometry in Data*, via intrinsic representations of shapes and data [27]. Spectral decomposition (Fig. 11) of shapes provides a new intrinsic framework for synthesizing complex shapes such as cerebral surfaces [35], and describing functions efficiently on these complex surfaces. This framework establishes the basics for machine learning of surface data [36]. An early application was conducted on retinotopy [57] (the study of functions in the visual cortex).

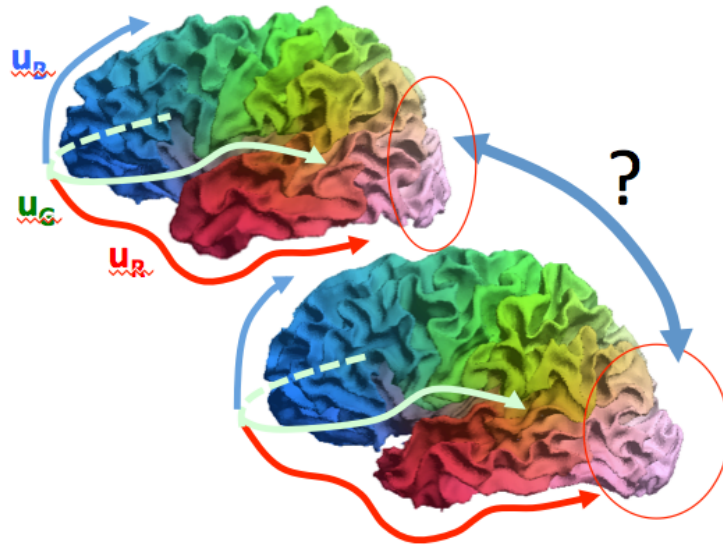


Figure 11. Spectral Representation – How to find an intrinsic representation of shapes for finding fast point correspondences and for learning surface data.

6.2.5. Consistency of the estimation of the template in quotient spaces

Participants: Loïc Devilliers [Correspondent], Stéphanie Allasonnière [Ecole Polytechnique], Xavier Penec.

Template estimation, Fréchet mean, quotient spaces

In [24], we studied the estimation of the template (the mean shape of our data) when the data is transformed by unknown group elements. In the case of a finite group acting isometrically on a linear space, we proved that the estimation of the template using the Fréchet mean in the quotient space is not always consistent.

6.3. Computational Physiology

6.3.1. Computational modeling of radiofrequency ablation for the planning and guidance of abdominal tumor treatment

Participants: Chloé Audigier [Correspondent], Hervé Delingette, Tommaso Mansi [Siemens], Nicholas Ayache.

This PhD work was carried out between the Asclepios research group, Inria Sophia Antipolis, France and Medical Imaging Technologies, Healthcare Technology Center, Siemens Medical Solutions USA, Princeton, NJ.

Radiofrequency Ablation Modeling, Patient-Specific Simulation, Lattice Boltzmann Method, Computer Model, Computational Fluid Dynamics, Heat Transfer, Cellular Necrosis, Parameter Estimation, Therapy Planning, Liver, Pre-clinical Study, Medical Imaging

Radio Frequency Ablation (RFA) is a minimally invasive therapy suited for liver tumor ablation. However, a patient-specific predictive tool is needed to plan and guide the treatment.

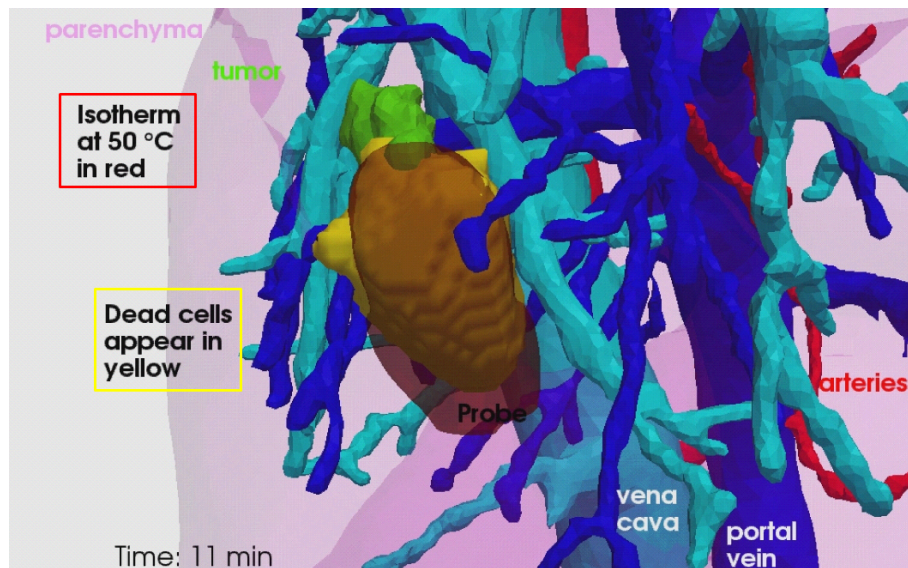


Figure 12. Computed isotherm at 50°C and computed necrosis in a subject-specific geometry.

We developed a computational framework for patient-specific planning of RFA, which includes the following contributions:

- A detailed computational model of the biophysical mechanisms (heat transfer, cellular necrosis, hepatic blood flow) involved in RFA of abdominal tumors based on patient images.
- A new implementation of the bio-heat equations coupled with a cellular necrosis model using the Lattice Boltzmann Method (LBM) on Graphics Processing Units (GPU), which allows near real-time computation.
- A Computational Fluid Dynamics (CFD) and porous media solver using LBM algorithm to compute the patient-specific blood flow in the hepatic circulatory system and the blood flow distribution inside the parenchyma.
- A complete patient-specific geometry including hepatic venous and arterial circulation system.
- The automatic estimation of the main parameters of the model. Two personalization strategies tested and evaluated on clinical and pre-clinical data.
- The evaluation of the proposed model on a clinical dataset of ten patients (see Fig. 12).
- The evaluation on a preclinical dataset of five swines from a comprehensive experimental set-up specially designed for RFA model validation.

The proposed RFA model and its evaluation on clinical data are presented in [10], and the evaluation of the RFA model on pre-clinical data is presented in [25]. The proposed model, its personalisation and its evaluation against clinical and preclinical data are presented in Chloé Audigier's PhD thesis [1].

6.3.2. Learning Cardiac Ablation Targets from Image Data and Simulation

Participants: Rocio Cabrera Lozoya [Correspondent], Maxime Sermesant, Nicholas Ayache.

This work was supported by the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Cardiac electrophysiology modeling, Intracardiac electrogram modeling, Machine learning, Radiofrequency ablation planning, electroanatomical mapping, local abnormal ventricular activities (LAVA)

Ventricular radiofrequency ablation can have a critical impact on preventing sudden cardiac arrest but it is challenging due to a highly complex arrhythmogenic substrate. We used advanced delayed enhanced-MR image characteristics in a machine learning framework to predict the presence of local abnormal ventricular activities (LAVA). Furthermore, we enriched these predictions through MR image-based patient-specific electrophysiology simulations and the modeling of normal and LAVA-like intracardiac electrograms using the dipole approach and their incorporation in the learning framework (see Fig. 13). Confidence maps can then be generated and analyzed prior to RFA to guide the intervention.

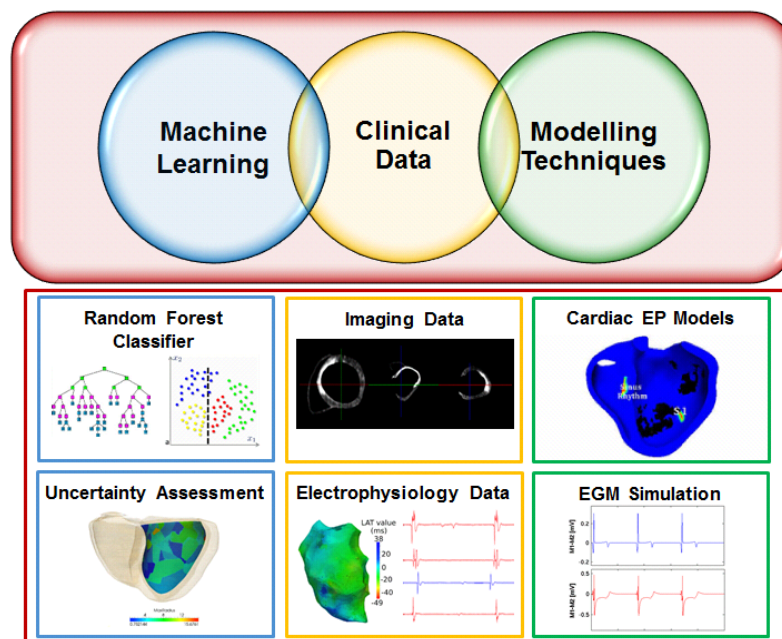


Figure 13. Coupled learning and simulation framework for LAVA identification.

6.3.3. Biophysical Modeling and Simulation of Longitudinal Brain MRIs with Atrophy in Alzheimer's Disease

Participants: Bishesh Khanal [Correspondent], Nicholas Ayache, Xavier Pennec.

This work has been partly supported by the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Alzheimer's Disease (AD), modeling brain deformation, biophysical model, simulation

- We developed a framework to generate patient specific multiple time-point images based on our biophysical model of brain deformation due to atrophy in Alzheimer's Disease (AD)[34]. From two time-point brain MRIs of a patient, we used the framework to simulate a new time-point brain MRI with the personalized atrophy for the patient (see Fig. 14).
- The framework can be used to evaluate methods that study the temporal relationships, ordering and co-evolution of atrophy in different structures of the brain.

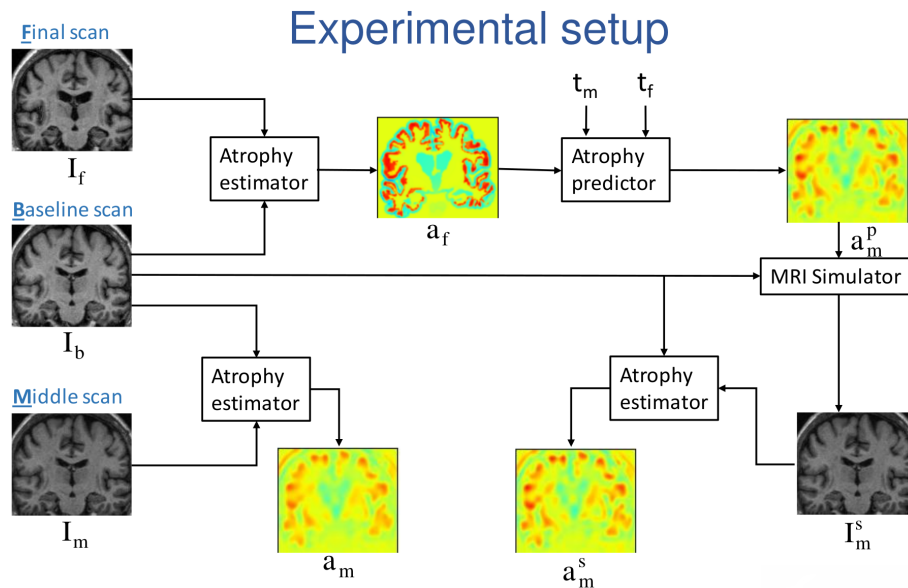


Figure 14. From a given baseline scan and a final scan at time t_f , a middle scan at time t_m is predicted and simulated. An experimental setup that allows comparing the atrophy (a_m^s) estimated from a simulated middle-scan against the atrophy (a_m) estimated from the real middle scan is shown.

6.3.4. Brain Tumor Growth Personalization and Segmentation Uncertainty

Participants: Matthieu Lê [Correspondent], Hervé Delingette, Jan Unkelbach, Nicholas Ayache.

This work is carried out between Asclepios research group, Inria Sophia Antipolis, France and the Department of Radiation Oncology of the Massachusetts General Hospital, Boston, USA. It is supported by the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Tumor growth, radiotherapy, modeling, personalization, segmentation, uncertainty, Bayesian

- We developed a method for the Bayesian personalization of a brain tumor growth model based on clinical MRIs [37] (see Fig. 15 Left).
- We proposed an algorithm for the sampling of several plausible segmentations, based on a single clinical segmentation (see Fig. 15 Right). This allows the uncertainty quantification of the radiotherapy plan based on several sample clinical target volumes [38]. This paper received the Young Scientist Award at the 2015 MICCAI conference in Munich, Germany.

6.3.5. Uncertainty quantification in personalised Cardiac models. Application to myocardial fiber uncertainty.

Participants: Roch-Philippe Molléro [Correspondent], Dominik Neumann [Siemens], Marc-Michel Rohé, Hervé Delingette, Maxime Sermesant, Xavier Pennec, Nicholas Ayache, Tommaso Mansi [Siemens].

This work was partly supported by the FP7 European project MD-Paedigree and was done in collaboration with Siemens Corporate Technology, Erlangen, Germany and Siemens Corporate Research, Princeton, New Jersey.

Heart Modeling - Myocardial Fibers - Biophysical Simulation - Uncertainty Quantification

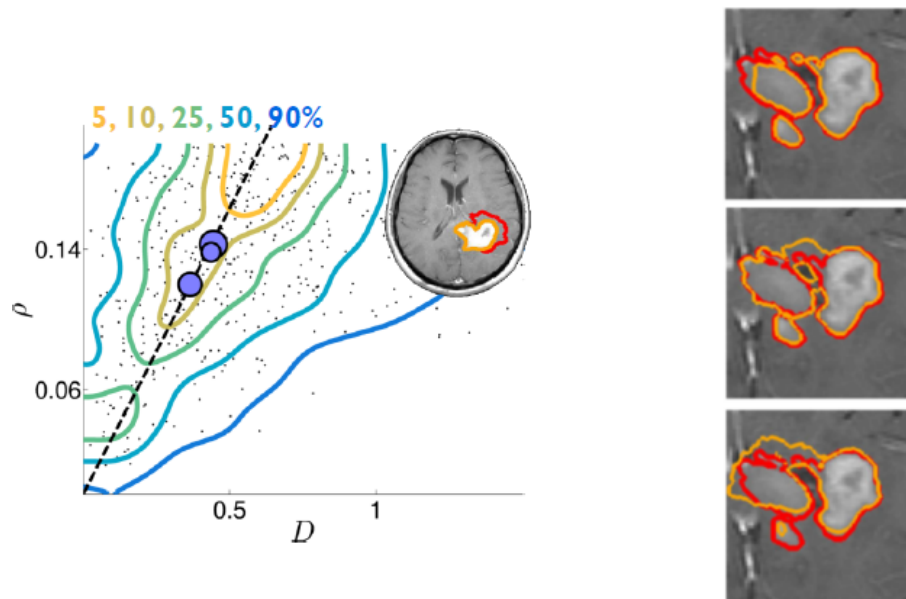


Figure 15. (Left) Bayesian personalization of a glioblastoma patient. Isocontours of the posterior probability of the diffusion parameter D and the proliferation parameter ρ ; (Right) Different sampled plausible segmentations in orange based on the clinical segmentation in red.

Computational models of the heart are of increasing interest for clinical applications due to their discriminative and predictive power. However, the personalisation step to go from a generic model to a patient-specific one is still challenging. In particular, it is still difficult to quantify the uncertainty on the estimated parameters and predicted values.

We developed a pipeline (see Fig. 16) to evaluate the impact of myocardial fibre uncertainty on the personalisation of an electromechanical model of the heart from ECG and medical images:

- We studied how to estimate the variability of the fibre architecture among a given population (from a myocardial fiber atlas).
- Then, we showed the variability of the personalised simulations, in electrophysiology (EP) and in biomechanics, with respect to the principal variations of the fibres.
- Finally, we discussed how the variations in this population of fibres impact the parameters of the personalised simulations.

This work led to a paper at FIMH 2015 conference in Maastricht, The Netherlands [45].

6.3.6. Non-invasive personalisation of the electrical heart model

Participants: Sophie Giffard-Roisin [Correspondent], Maxime Sermesant, Nicholas Ayache, Hervé Delingette.

This work has been supported by the European Project FP7 under grant agreement VP2HF (no 611823) and the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Cardiac Modelling, Personalised Simulation, Electrical Simulation

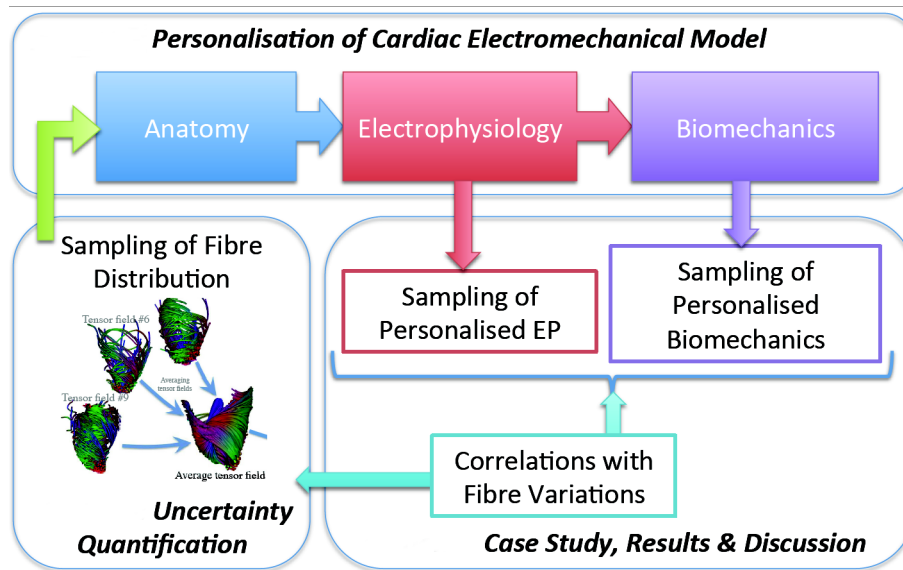


Figure 16. Global scheme of fibre variability propagation along the personalisation pipeline.

Non-invasive cardiac electrical data has been acquired at St Thomas' Hospital, London. It consists in Body Surface Potential Mapping (BSPM), which are recordings of the electrical potential on several locations on the surface of the torso (see Fig. 17). From BSPMs and MRI data of the heart, we aim at personalizing the electrical propagation model of the heart previously developed within the Asclepios team.

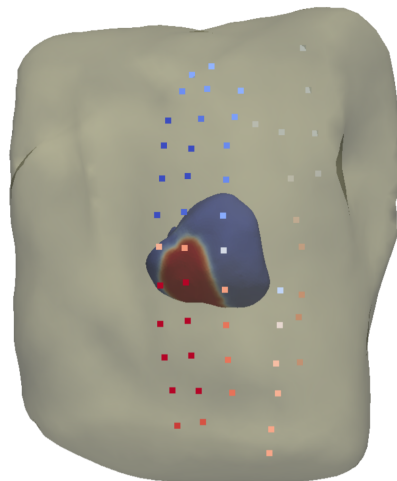


Figure 17. Torso representation used for personalizing cardiac electrical parameters from non-invasive observations.

ATHENA Project-Team

6. New Results

6.1. Modeling in Diffusion MRI

6.1.1. *Improving fiber alignment in HARDI by combining contextual PDE flow with constrained spherical deconvolution*

Participants: Jorg Portegies [Department of Mathematics and Computer Science, Eindhoven University of Technology], Rutger Fick, Gonzalo Sanguinetti [Department of Mathematics and Computer Science, Eindhoven University of Technology], Shephan Meesters [Department of Mathematics and Computer Science, Eindhoven University of Technology], Gabriel Girard [Athena, Inria Sophia-A-M & SCIL Lab., Sherbrooke University], Remco Duits [Department of Mathematics and Computer Science, Eindhoven University of Technology].

We propose two strategies to improve the quality of tractography results computed from diffusion weighted magnetic resonance imaging (DW-MRI) data. Both methods are based on the same PDE framework, defined in the coupled space of positions and orientations, associated with a stochastic process describing the enhancement of elongated structures while preserving crossing structures. In the first method we use the enhancement PDE for contextual regularization of a fiber orientation distribution (FOD) that is obtained on individual voxels from high angular resolution diffusion imaging (HARDI) data via constrained spherical deconvolution (CSD). Thereby we improve the FOD as input for subsequent tractography. Secondly, we introduce the fiber to bundle coherence (FBC), a measure for quantification of fiber alignment. The FBC is computed from a tractography result using the same PDE framework and provides a criterion for removing the spurious fibers. We validate the proposed combination of CSD and enhancement on phantom data and on human data, acquired with different scanning protocols. On the phantom data we find that PDE enhancements improve both local metrics and global metrics of tractography results, compared to CSD without enhancements. On the human data we show that the enhancements allow for a better reconstruction of crossing fiber bundles and they reduce the variability of the tractography output with respect to the acquisition parameters. Finally, we show that both the enhancement of the FODs and the use of the FBC measure on the tractography improve the stability with respect to different stochastic realizations of probabilistic tractography. This is shown in a clinical application: the reconstruction of the optic radiation for epilepsy surgery planning.

This work has been published in [19]

6.1.2. *Sparse reconstruction challenge for diffusion MRI: Validation on a physical phantom to determine which acquisition scheme and analysis method to use?*

Participants: Lipeng Ning [Brigham and Women's Hospital, Harvard Medical School, Boston], Frederik Laun [German Cancer Research Institute], Yogesh Rathi [Brigham and Women's Hospital, Harvard Medical School, Boston], Thinhinane Megherbi [ParIMed Team, LRPE, USTHB, Algiers], Mario Zuccheli [Dpt of Computer Science, University of Verona], Gloria Menegaz [Dpt of Computer Science, University of Verona], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Aurobrata Ghosh, Rutger Fick, Rachid Deriche.

Diffusion magnetic resonance imaging (dMRI) is the modality of choice for investigating in-vivo white matter connectivity and neural tissue architecture of the brain. The diffusion-weighted signal in dMRI reflects the diffusivity of water molecules in brain tissue and can be utilized to produce image-based biomarkers for clinical research. Due to the constraints on scanning time, a limited number of measurements can be acquired within a clinically feasible scan time. In order to reconstruct the dMRI signal from a discrete set of measurements, a large number of algorithms have been proposed in recent years in conjunction with varying sampling schemes, i.e., with varying b-values and gradient directions. Thus, it is imperative to compare the performance of these reconstruction methods on a single data set to provide appropriate guidelines to

neuroscientists on making an informed decision while designing their acquisition protocols. For this purpose, the SPArse Reconstruction Challenge (SPARC) was held along with the workshop on Computational Diffusion MRI (at MICCAI 2014) to validate the performance of multiple reconstruction methods using data acquired from a physical phantom. A total of 16 reconstruction algorithms (9 teams) participated in this community challenge. The goal was to reconstruct single b-value and/or multiple b-value data from a sparse set of measurements. In particular, the aim was to determine an appropriate acquisition protocol (in terms of the number of measurements, b-values) and the analysis method to use for a neuroimaging study. The challenge did not delve on the accuracy of these methods in estimating model specific measures such as fractional anisotropy (FA) or mean diffusivity, but on the accuracy of these methods to fit the data. This work presents several quantitative results pertaining to each reconstruction algorithm. The conclusions in this work provide a valuable guideline for choosing a suitable algorithm and the corresponding data-sampling scheme for clinical neuroscience applications.

This work has been published in [18].

6.1.3. *A Unifying framework for spatial and temporal diffusion in dMRI*

Participants: Rutger Fick, Demian Wassermann, Marco Pizzolato, Rachid Deriche.

We propose a novel framework to simultaneously represent the diffusion-weighted MRI (dMRI) signal over diffusion times, gradient strengths and gradient directions. Current frameworks such as the 3D Simple Harmonic Oscillator Reconstruction and Estimation basis (3D-SHORE) only represent the signal over the spatial domain, leaving the temporal dependency as a fixed parameter. However, microstructure- focused techniques such as Axcaliber and ActiveAx provide evidence of the importance of sampling the dMRI space over diffusion time. Up to now there exists no generalized framework that simultaneously models the dependence of the dMRI signal in space and time. We use a functional basis to fit the 3D+t spatio-temporal dMRI signal, similarly to the 3D-SHORE basis in three dimensional 'q-space'. The lowest order term in this expansion contains an isotropic diffusion tensor that characterizes the Gaussian displacement distribution, multiplied by a negative exponential. We regularize the signal fitting by minimizing the norm of the analytic Laplacian of the basis. The continuous 3D+t signal representation can provide new insights on the anomalous nature of the dMRI signal in human tissues, i.e., when mean-squared molecular displacements varies slower than linearly with the diffusion time. From the fitting one can also estimate the axon radius distribution parameters along any direction using approaches similar to AxCaliber. We validate our technique on synthetic data generated using the theoretical model proposed by Callaghan et al. We show that our method is robust to noise and can accurately describe the restricted spatio-temporal signal decay originating from tissue models such as cylindrical pores. Moreover, we apply our method on real data from an ActiveAx acquisition. Overall our approach allows to represent the complete 3D+t dMRI signal which should prove helpful in understanding normal and pathologic nervous tissue.

This work has been published in [26]

6.1.4. *Exploiting the phase in dMRI for microstructure recovery: Towards axonal tortuosity via asymmetric diffusion processes*

Participants: Marco Pizzolato, Demian Wassermann, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Microstructure recovery procedures via Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) usually discard the signal's phase, assuming symmetry in the underlying diffusion process. In this work, we propose to recover the Ensemble Average Propagator (EAP) directly from the complex DW signal in order to describe also eventual diffusional asymmetry, thus obtaining an asymmetric EAP. The asymmetry of the EAP is then related to tortuosity of undulated white matter axons, which are found in pathological scenarios associated with axonal elongation or compression. We derive a model of the EAP for this geometry and quantify its asymmetry. Results show that the EAP obtained when accounting for the DW signal's phase provides useful microstructural information in such pathological scenarios. Furthermore, we validate these results in-silico through 3D Monte-Carlo simulations of white matter tissue that has experienced different degrees of elongation/compression.

This work has been published in [35]

6.1.5. *A temperature phantom to probe the Ensemble Average Propagator asymmetry: an in-silico study*

Participants: Marco Pizzolato, Demian Wassermann, Tanguy Duval [Institute of Biomedical Engineering, Polytechnique Montréal, Montréal], Jennifer Campbell [Montreal Neurological Institute, McGill University], Timothé Boutelier [Olea Medical, La Ciotat], Julien Cohen-Adad [Institute of Biomedical Engineering, Polytechnique Montréal, Montréal], Rachid Deriche.

The detection and quantification of asymmetry in the Ensemble Average Propagator (EAP) obtained from the Diffusion-Weighted (DW) signal has been shown only for theoretical models. EAP asymmetry appears for instance when diffusion occurs within fibers with particular geometries. However the quantification of EAP asymmetry corresponding to such geometries in controlled experimental conditions is limited by the difficulty of designing fiber geometries on a micrometer scale. To overcome this limitation we propose to adopt an alternative paradigm to induce asymmetry in the EAP. We apply a temperature gradient to a spinal cord tract to induce a corresponding diffusivity profile that alters locally the diffusion process to be asymmetric. We simulate the EAP and the corresponding complex DW signal in such a scenario. We quantify EAP asymmetry and investigate its relationship with the applied experimental conditions and with the acquisition parameters of a Pulsed Gradient Spin-Echo sequence. Results show that EAP asymmetry is sensible to the applied temperature-induced diffusivity gradient and that its quantification is influenced by the selected acquisition parameters.

This work has been published in [36]

6.1.6. *How to get more out of a clinically feasible 64 gradient dMRI acquisition: multi-shell versus single-shell*

Participants: Rutger Fick, Mario Zuccheli [Dpt of Computer Science, University of Verona], Gabriel Girard [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Gloria Menegaz [Dpt of Computer Science, University of Verona], Rachid Deriche.

For clinical applications the number of diffusion MRI (dMRI) samples that can be obtained is often limited by scanner time and patient comfort. For this reason one often uses short scanning protocols that acquire just 32 or 64 gradient directions using a single b-value to obtain diffusion measures such as the fractional anisotropy from Diffusion Tensor Imaging (DTI) or to estimate the white matter orientation using Constrained Spherical Deconvolution (CSD). Using 3D-SHORE and MAP-MRI, we show that by spreading the same number of dMRI samples over different b-shells (sampling angularly and radially) we can estimate not only the directionality of the white matter using the ODF, but also the radially dependent higher order diffusion measures that SHORE and MAP-MRI provide. This approach lends itself well for situations where acquisition time is limited, and is therefore particularly well suited for clinical applications.

This work has been published in [29].

6.2. Tissue Microstructures features recovery & applications

6.2.1. *Laplacian-regularized MAP-MRI : Improving axonal caliber estimation*

Participants: Rutger Fick, Demian Wassermann, Gonzalo Sanguinetti, Rachid Deriche.

In diffusion MRI, the accurate description of the entire diffusion signal from sparse measurements is essential to enable the recovery of microstructural information of the white matter. The recent Mean Apparent Propagator (MAP)-MRI basis is especially well suited for this task, but the basis fitting becomes unreliable in the presence of noise. As a solution we propose a fast and robust analytic Laplacian regularization for MAP-MRI. Using both synthetic diffusion data and human data from the Human Connectome Project we show that (1) MAP-MRI has more accurate microstructure recovery compared to classical techniques, (2) regularized MAP-MRI has lower signal fitting errors compared to the unregularized approach and a positivity constraint on the EAP and (3) that our regularization improves axon radius recovery on human data.

This work has been published in [27]

6.2.2. *Using 3D-SHORE and MAP-MRI to obtain both tractography and microstructural contrasts from a clinical dMRI acquisition*

Participants: Rutger Fick, Mario Zuccheli [Dpt of Computer Science, University of Verona], Gabriel Girard [Athena, Inria Sophia-A-M & SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Gloria Menegaz [Dpt of Computer Science, University of Verona], Rachid Deriche.

Diffusion MRI (dMRI) is used to characterize the directionality and microstructural properties of brain white matter (WM) by measuring the diffusivity of water molecules. In clinical practice the number of dMRI samples that can be obtained is limited, and one often uses short scanning protocols that acquire just 32 to 64 different gradient directions using a single gradient strength (b-value). Such 'single shell' scanning protocols restrict one to use methods that have assumptions on the radial decay of the dMRI signal over different b-values, which introduces estimation biases. In this work, we show, that by simply spreading the same number of samples over multiple b-values (i.e. multi-shell) we can accurately estimate both the WM directionality using 3D-SHORE and characterize the radially dependent diffusion microstructure measures using MAP-MRI. We validate our approach by undersampling both noisy synthetic and human brain data of the Human Connectome Project, proving this approach is well-suited for clinical applications.

This work has been published in [28]

6.2.3. *A sensitivity analysis of Q-space Indices with respect to changes in axonal diameter, dispersion and tissue composition*

Participants: Rutger Fick, Marco Pizzolato, Demian Wassermann, Mario Zuccheli [Dpt of Computer Science, University of Verona], Gloria Menegaz [Dpt of Computer Science, University of Verona], Rachid Deriche.

In Diffusion MRI, q-space indices are scalar quantities that describe properties of the ensemble average propagator (EAP). Their values are often linked to the axonal diameter – assuming that the diffusion signal originates from inside an ensemble of parallel cylinders. However, histological studies show that these assumptions are incorrect, and axonal tissue is often dispersed with various tissue compositions. Direct interpretation of these q-space indices in terms of tissue change is therefore impossible, and we must treat them as scalars that only give non-specific contrast – just as DTI indices. In this work, we analyze the sensitivity of q-space indices to tissue structure changes by simulating axonal tissue with changing axonal diameter, dispersion and tissue compositions. Using human connectome project data we then predict which indices are most sensitive to tissue changes in the brain. We show that, in both multi-shell and single-shell (DTI) data, q-space indices have higher sensitivity to tissue changes than DTI indices in large parts of the brain. Based on these results, it may be interesting to revisit older DTI studies using q-space indices as a marker for pathology.

This work has been accepted at the conference ISBI 2016.

6.2.4. *MAPL: Tissue microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP data*

Participants: Rutger Fick, Demian Wassermann, Emanuel Caruyer, Rachid Deriche.

The recovery of microstructure-related features of the brain's white matter is a current challenge in diffusion MRI. To robustly estimate these important features from diffusion MRI data, we propose to analytically regularize MAP-MRI's coefficient estimation using the norm of the Laplacian of the reconstructed signal. We first compare our approach, which we call MAPL, with competing state-of-the-art functional basis approaches. We show that it outperforms the original MAP-MRI implementation and the recently proposed modified Spherical Polar Fourier (mSPF) basis with respect to signal fitting, EAP and ODF reconstruction in noisy, sparsely sampled data of a physical phantom with reference gold standard data. Then, to reduce the variance of parameter estimation using multi-compartment tissue models, we propose to use MAPL's signal

fitting and extrapolation as a preprocessing step. We study the effect of MAPL on the estimation of axon diameter using a simplified Axcaliber model and axonal dispersion using the Neurite Orientation Dispersion and Density Imaging (NODDI) model. We show the positive effect of using it as a preprocessing step in estimating and reducing the variances of these parameters in the Corpus Callosum of six different subjects of the MGH Human Connectome Project. Finally we correlate the estimated axon diameter, dispersion and restricted volume fractions with Fractional Anisotropy (FA) and clearly show that changes in FA significantly correlate with changes with all estimated parameters. Overall, we illustrate the potential of using a well-regularized functional basis together with multi-compartment approaches to recover important microstructure tissue parameters with much less variability, thus contributing to the challenge of better understanding microstructure-related features of the brain's white matter.

This work has been submitted to the journal *NeuroImage*.

6.3. Towards microstructural based tractography

6.3.1. *AxTract: Microstructure-driven tractography based on the Ensemble Average Propagator*

Participants: Gabriel Girard [Athena, Inria Sophia-A-M & SCIL Lab., Sherbrooke University], Rutger Fick, Maxime Descoteaux [SCIL Lab., Sherbrooke University], Demian Wassermann, Rachid Deriche.

In this work, we propose a novel method to simultaneously trace brain white matter (WM) fascicles and estimate WM microstructure characteristics. Recent advancements in diffusion-weighted imaging (DWI) allow multi-shell acquisitions with b-values of up to 10,000 s/mm² in human subjects, enabling the measurement of the ensemble average propagator (EAP) at distances as short as 10 micro-meters. Coupled with continuous models of the full 3D DWI signal and the EAP such as Mean Apparent Propagator (MAP) MRI, these acquisition schemes provide unparalleled means to probe the WM tissue in vivo. Presently, there are two complementary limitations in tractography and microstructure measurement techniques. Tractography techniques are based on models of the DWI signal geometry without taking specific hypotheses of the WM structure. This hinders the tracing of fascicles through certain WM areas with complex organization such as branching, crossing, merging, and bottlenecks that are indistinguishable using the orientation-only part of the DWI signal. Microstructure measuring techniques, such as AxCaliber, require the direction of the axons within the probed tissue before the acquisition as well as the tissue to be highly organized. Our contributions are twofold. First, we extend the theoretical DWI models proposed by Callaghan et al. to characterize the distribution of axonal calibers within the probed tissue taking advantage of the MAP-MRI model. Second, we develop a simultaneous tractography and axonal caliber distribution algorithm based on the hypothesis that axonal caliber distribution varies smoothly along a WM fascicle. To validate our model we test it on in-silico phantoms and on the HCP dataset

This work has been published in [23]

6.3.2. *Studying white matter tractography reproducibility through connectivity matrices*

Participants: Gabriel Girard [Athena, Inria Sophia-A-M & SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion-weighted imaging is often used as a starting point for in vivo white matter (WM) connectivity to reconstruct potential WM pathways between brain areas. In this study, we investigate the reproducibility of the connectivity matrix, resulting from different tractography parameters. We vary the number of streamlines used to construct the matrix in cortical to cortical connectivity and analyze its effects. We also compare the effect of probabilistic and deterministic local streamline tractography algorithms, seeding both from the WM and from WM-grey matter interface.

This work has been published in [31]

6.3.3. *Structural connectivity reproducibility through multiple acquisitions*

Participants: Gabriel Girard [Athena, Inria Sophia-A-M & SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

dMRI is often used to reconstruct white matter pathways between brain areas for in vivo brain connectivity. In this study, we investigate the reproducibility and the specificity of connectivity matrices in cortico-cortical connectivity using probabilistic and deterministic streamline tractography, seeding from both the white matter and the white matter-grey matter interface.

This work has been published in [30]

6.4. Computational Diffusion MRI

6.4.1. *Robust and efficient linear registration of white-matter fascicles in the space of streamlines*

Participants: Eleftherios Garyfallidis [SCIL Lab., Sherbrooke University], Omar Cepeda [SCIL Lab., Sherbrooke University], Demian Wassermann, Maxime Descoteaux [SCIL Lab., Sherbrooke University].

The neuroscientific community is very much interested in analyzing specific white matter bundles like the arcuate fasciculus, the corticospinal tract, or the recently discovered Aslant tract to study sex differences, lateralization and many other connectivity applications. For this reason, experts spend time manually segmenting these fascicles and bundles using streamlines obtained from diffusion MRI tractography. However, to date, there are very few computational tools available to register these fascicles directly so that they can be analyzed and their differences quantified across populations. In this work, we introduce a novel, robust and efficient framework to align bundles of streamlines directly in the space of streamlines. We call this framework Streamline-based Linear Registration. We first show that this method can be used successfully to align individual bundles as well as whole brain streamlines. Additionally, if used as a piecewise linear registration across many bundles, we show that our novel method systematically provides higher overlap (Jaccard indices) than state-of-the-art nonlinear image-based registration in the white matter. We also show how our novel method can be used to create bundle-specific atlases in a straightforward manner and we give an example of a probabilistic atlas construction of the optic radiation. In summary, Streamline-based Linear Registration provides a solid registration framework for creating new methods to study the white matter and perform group-level tractometry analysis.

This work has been published in [14]

6.4.2. *Cortical surface parcellation via dMRI using mutual nearest neighbor condition*

Participants: Brahim Belaoucha, Maureen Clerc, Théodore Papadopoulos.

In this work, we present a method that aims at parcellating the cortical surface from individual anatomy. The parcellation is obtained using the mutual nearest neighbor criteria to obtain regions that have similar fiber distribution. The later is obtained by applying a probabilistic tractography on the diffusion MRI (dMRI), a non-invasive modality allowing the access to the structural information of the cortical surface. The proposed algorithm is compared to some of the atlases that can be found in the literature. We show that these atlases have lower similarity of fibers distributions than the proposed algorithm.

This work has been accepted at the conference ISBI 2016.

6.5. Clinical and Neurocognitive Applications of Diffusion MRI

6.5.1. *Plasticity of left perisylvian white-matter tracts is associated with individual differences in math learning brain structure and function*

Participants: Dietsje Jolles [Stanford University & Leiden University], Demian Wassermann, Ritika Chokhani [Stanford University], Jennifer Richardson [Stanford University], Caitlin Tenison [Stanford University], Roland Bammer [Stanford University], Lynn Fuchs [Vanderbilt University], Kaustubh Supekar [Stanford University], Vinod Menon [Stanford University].

Plasticity of white matter tracts is thought to be essential for cognitive development and academic skill acquisition in children. However, a dearth of high-quality diffusion tensor imaging (DTI) data measuring longitudinal changes with learning, as well as methodological difficulties in multi-time point tract identification have limited our ability to investigate plasticity of specific white matter tracts. Here, we examine learning-related changes of white matter tracts innervating inferior parietal, prefrontal and temporal regions following an intense two-month math tutoring program. DTI data were acquired from 18 third grade children, both before and after tutoring. A novel fiber tracking algorithm based on a White Matter Query Language (WMQL) was used to identify three sections of the superior longitudinal fasciculus (SLF) linking frontal and parietal (SLF-FP), parietal and temporal (SLF-PT) and frontal and temporal (SLF-FT) cortices, from which we created child-specific probabilistic maps. The SLF-FP, SLF-FT, and SLF-PT tracts identified with the WMQL method were highly reliable across the two time points and showed close correspondence to tracts previously described in adults. Notably, individual differences in behavioral gains after two months of tutoring were specifically correlated with plasticity in the left SLF-FT tract. Our results extend previous findings of individual differences in white matter integrity, and provide important new insights into white matter plasticity related to math learning in childhood. More generally, our quantitative approach will be useful for future studies examining longitudinal changes in white matter integrity associated with cognitive skill development.

This work has been published in [16]

6.5.2. Prefrontal cortex white matter tracts in prodromal Huntington disease

Participants: Joy T. Matsui [Iowa University], Jatin G. Vaidya [Iowa University], Demian Wassermann [Iowa University], Regina Eunyong Kim [Iowa University], Vincent A. Magnotta [Iowa University], Hans J. Johnson [Iowa University], Jane S. Paulsen [Iowa University], Predict-Hd Investigators And Coordinators Of The Huntington Study Group [NIH].

Huntington disease (HD) is most widely known for its selective degeneration of striatal neurons but there is also growing evidence for white matter (WM) deterioration. The primary objective of this research was to conduct a large-scale analysis using multi-site diffusion-weighted imaging (DWI) tractography data to quantify diffusivity properties along major prefrontal cortex WM tracts in prodromal HD. Fifteen international sites participating in the PREDICT-HD study collected imaging and neuropsychological data on gene-positive HD subjects without a clinical diagnosis (i.e. prodromal) and gene-negative control subjects. The anatomical prefrontal WM tracts of the corpus callosum (PFCC), anterior thalamic radiations (ATR), inferior fronto-occipital fasciculi (IFO), and uncinate fasciculi (UNC) were identified using streamline tractography of DWI. Within each of these tracts, tensor scalars for fractional anisotropy, mean diffusivity, radial diffusivity, and axial diffusivity coefficients were calculated. We divided prodromal HD subjects into three CAG-age product (CAP) groups having Low, Medium, or High probabilities of onset indexed by genetic exposure. We observed significant differences in WM properties for each of the four anatomical tracts for the High CAP group in comparison to controls. Additionally, the Medium CAP group presented differences in the ATR and IFO in comparison to controls. Furthermore, WM alterations in the PFCC, ATR, and IFO showed robust associations with neuropsychological measures of executive functioning. These results suggest that long-range tracts essential for cross-region information transfer show early vulnerability in HD and may explain cognitive problems often present in the prodromal stage.

This work has been published in [17]

6.6. Perfusion MRI

6.6.1. Perfusion MRI deconvolution with delay estimation and non-negativity constraints

Participants: Marco Pizzolato, Auro Ghosh, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Perfusion MRI deconvolution aims to recover the time-dependent residual amount of indicator (residue function) from the measured arterial and tissue concentration time-curves. The deconvolution is complicated by the presence of a time lag between the measured concentrations. Moreover the residue function must be non-negative and its shape may become non-monotonic due to dispersion phenomena. We introduce Modified Exponential Bases (MEB) to perform deconvolution. The MEB generalizes the previously proposed exponential approximation (EA) by taking into account the time lag and introducing non-negativity constraints for the recovered residue function also in the case of non-monotonic dispersed shapes, thus overcoming the limitation due to the non-increasing assumption of the EA. The deconvolution problem is solved linearly. Quantitative comparisons with the widespread block-circulant Singular Value Decomposition show favorable results in recovering the residue function.

This work has been published in [34]

6.6.2. *Elucidating dispersion effects in perfusion MRI by means of dispersion-compliant bases*

Participants: Marco Pizzolato, Rutger Fick, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Dispersion effects in perfusion MRI data have a relevant influence on the residue function computed from deconvolution of the measured arterial and tissular concentration time-curves. Their characterization allows reliable estimation of hemodynamic parameters and can reveal pathological tissue conditions. However the time-delay between the measured concentration time-curves is a confounding factor. We perform deconvolution by means of dispersion-compliant bases, separating dispersion from delay effects. In order to characterize dispersion we introduce shape parameters, such as the dispersion time and index. We propose a new formulation for the dispersed residue function and perform in-silico experiments that validate the reliability of our approach against the block-circulant Singular Value Decomposition. We successfully apply the approach to stroke MRI data and show that the calculated parameters are coherent with physiological considerations, highlighting the importance of dispersion as an effect to be measured rather than discarded.

This work has been accepted at the conference ISBI 2016.

6.6.3. *Unveiling the dispersion kernel in DSC-MRI by means of dispersion-compliant bases and control point interpolation techniques*

Participants: Marco Pizzolato, Rutger Fick, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

In DSC-MRI the presence of dispersion affects the estimation, via deconvolution, of the residue function that characterizes the perfusion in each voxel. Dispersion is described by a Vascular Transport Function (VTF) which knowledge is essential to recover a dispersion-free residue function. State-of-the-art techniques aim at characterizing the VTF but assume a specific shape for it, which in reality is unknown. We propose to estimate the residue function without assumptions by means of Dispersion-Compliant Bases (DCB). We use these results to find which VTF model better describes the in-vivo data for each tissue type by means of control point interpolation approaches.

This work has been submitted to the conference ISMRM 2016.

6.6.4. *Improved vascular transport function characterization in DSC-MRI via deconvolution with dispersion-compliant bases*

Participants: Marco Pizzolato, Rutger Fick, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Bolus dispersion phenomena affect the residue function computed via deconvolution of DSC-MRI data. Indeed the obtained effective residue function can be expressed as the convolution of the true one with a Vascular Transport Function (VTF) that characterizes the dispersion. The state-of-the-art technique CPI+VTF allows to estimate the actual residue function by assuming a model for the VTF. We propose to perform deconvolution representing the effective residue function with Dispersion-Compliant Bases (DCB) without assumptions on the VTF, and then apply the CPI+VTF on DCB results. We show that DCB improve robustness to noise and allow to better characterize the VTF.

This work has been submitted to the conference ISMRM 2016.

6.7. MEG, EEG and cochlear modeling

6.7.1. MEM-diffusion MRI framework to solve MEEG inverse problem

Participants: Brahim Belaoucha, Jean-Marc Lina, Maureen Clerc, Théodore Papadopoulos.

In this work, we present a framework to fuse information coming from diffusion magnetic resonance imaging (dMRI) with Magnetoencephalography (MEG)/ Electroencephalography (EEG) measurements to reconstruct the activation on the cortical surface. The MEG/EEG inverse-problem is solved by the Maximum Entropy on the Mean (MEM) principle and by assuming that the sources inside each cortical region follow Normal distribution. These regions are obtained using dMRI and assumed to be functionally independent. The source reconstruction framework presented in this work is tested using synthetic and real data. The activated regions for the real data is consistent with the literature about the face recognition and processing network.

This work was published in the proceedings of the conference EUSIPCO 2015 [22].

6.7.2. MEG/EEG reconstruction in the reduced source space

Participants: Brahim Belaoucha, Théodore Papadopoulos.

Obtaining the brain activity with the distributed source model from MEG or EEG measurements is ill-posed problem due to the high number of unknowns compared to the number of measurements. The idea of this work is to reduce the solution space size from the number of sources to a smaller space. Assuming that sources inside each functional region have equal activation allows us to reduce the number of columns in the leadfield matrix from the number of nodes S required to model the cortex to a number of regions K , which is much smaller. These regions are obtained from a dMRI parcellation-based region growing algorithm. A region is assumed to contain sources that have similar fibers distribution. To obtain a sparse solution, we assume that only a few regions are active simultaneously. BIC1 is used to obtain the optimal number of regions (K_p) that explains the MEG/EEG data.

We compared the results of the proposed method to the ones from Minimum Norm Estimate (MNE) and LASSO. The first gives a smooth solution and the second gives a sparse solution. To test the accuracy of the reconstruction, we activated simultaneously from two to five regions in both hemispheres with synthetic low SNR signals (10 dB). Our approach could detect the right number of activated regions and provided more accurate reconstructions compared to MNE and LASSO.

Our approach assumes that few regions are active simultaneously which allows us to reduce the space to a few unknowns. It can be seen as an approximation to the l_0 norm. Even though assuming a constant activation in each functional region is a hard constraint, it allows us to reduce the space size from S to K . The obtained solution can be used to detect extended sources (e.g epileptic activity) or as an initialization step to other approaches to obtain more detailed solutions in the active regions.

This work was presented at the conference BaCI 2015 [24].

6.7.3. Realistic simulation of electric potential distributions of different stimulation modes in an implanted cochlea

Participants: Kai Dang, Maureen Clerc, Clair Vandersteen [Institut Universitaire de la Face et du Cou, Nice], Nicolas Guevara [Institut Universitaire de la Face et du Cou, Nice], Dan Gnansia [Oticon Medical/Neurelec].

Simulation of the intracochlear potentials is an important approach to study the activation of auditory nerve fibers under electrical stimulations. However, it is still unclear to which extent the simulation results are affected by precision in reproducing the exact cochlear geometry. In this study, we address to this question by comparing the actual electric potential measured from implanted human specimen with the simulation outputs from two different parametric 3D cochlear models. One of the models is created from the default values while the other is adapted to the micro-CT scan data of the implanted cochlea.

This work was presented at the Association of Research in Otolaryngology 38th MidWinter Meeting, Feb 2015, Baltimore, United States [38].

We also made an in situ validation of electrical models: Cochlear implants have been proved to be an effective treatment for patients with sensorineural hearing loss. Among all the approaches that have been developed to design better cochlear implants, 3D model-based simulation stands out due to its detailed description of the electric field which helps reveal the electrophysiological phenomena inside the cochlea. With the advances in the cochlear implant manufacturing technology, the requirement on simulation accuracy increases. Improving the simulation accuracy relies on two aspects: 1) a better geometrical description of the cochlea that is able to distinguish the subtle differences across patients; 2) a comprehensive and reliable validation of the created 3D model. In this paper, targeting at high precision simulation, we propose a parametric cochlea model which uses micro-CT images to adapt to different cochlea geometries, then demonstrate its validation process with multi-channel stimulation data measured from a implanted cochlea. Comparisons between the simulation and validation data show a good match under a variety of stimulation configurations. The results suggest that the electric field distribution is affected by the geometric characteristics of each individual cochlea. These differences can be correctly reflected by simulations based on a 3D model tuned with personalized data.

This work was presented at the 7th International IEEE EMBS Conference on Neural Engineering, Apr 2015, Montpellier, France [25].

6.7.4. Influence of skull modelling on conductivity estimation for EEG source analysis

Participants: Christos Papageorgakis, Maureen Clerc, Benjamin Lanfer [BESA GmbH].

The skull conductivity strongly influences the accuracy of EEG source localization methods. As the conductivity of the skull has strong inter-individual variability, conductivity estimation techniques are required. Typically, conductivity estimation is performed on data from a single event-related stimulation paradigm, which can be explained by one dipole source. A conductivity value for the skull can be estimated as the value for which the single dipole source provides the best goodness of fit to the data. This conductivity value is then used to analyse the actual data of interest. It is known that the optimal local skull conductivity when modelling the skull as one compartment depends on the amount of spongiosa present locally. The research question arising is: Is conductivity estimation based on data from a single paradigm meaningful without accounting for the internal skull structure ?

This work was presented at the conference BaCI 2015 [33], and is submitted for journal publication.

6.7.5. Dictionary learning for M/EEG multidimensional data

Participants: Christos Papageorgakis, Sebastian Hitziger, Théodore Papadopoulo.

Signals obtained from magneto- or electroencephalography (M/EEG) are very noisy and inherently multidimensional, i.e. provide a vector of measurements at each single time instant. To cope with noise, researchers traditionally acquire measurements over multiple repetitions (trials) and average them to classify various patterns of activity. This is not optimal because of trial-to-trial variability (waveform variation, jitters). The jitter-adaptive dictionary learning method (JADL) has been developed to better handle for this variability (with a particular emphasis on jitters). JADL is a data-driven method that learns a dictionary (prototype pieces) from a set of signals, but is currently limited to a single channel, which restricts its capacity to work with very noisy data such as M/EEG. We propose an extension to the jitter-adaptive dictionary learning method, that is able to handle multidimensional measurements such as M/EEG.

This work was presented at the conference BaCI 2015 [32].

6.8. Brain Computer Interfaces

6.8.1. Decoding covert shifts of attention induced by ambiguous visuospatial cues

Participants: Romain Trachel, Maureen Clerc, Thomas Brochier [Institut de Neurosciences de la Timone, Marseille].

Simple and unambiguous visual cues (e.g., an arrow) can be used to trigger covert shifts of visual attention away from the center of gaze. The processing of visual stimuli is enhanced at the attended location. Covert shifts of attention modulate the power of cerebral oscillations in the alpha band over parietal and occipital regions. These modulations are sufficiently robust to be decoded on a single trial basis from electroencephalography (EEG) signals. It is often assumed that covert attention shifts are under voluntary control, and that they also occur in more natural and complex environments, but there is no direct evidence to support this assumption. We address this important issue by using random-dot stimuli to cue one of two opposite locations, where a visual target is presented. We contrast two conditions, one in which the random-dot motion is predictive of the target location, and the other, in which it provides ambiguous information. Behavioral results show attention shifts in anticipation of the visual target, in both conditions. In addition, using the common spatial patterns (CSPs) algorithm, we extract EEG power features in the alpha-band (around 10 Hz) that best discriminate the attended location in single trials. We obtain a significant decoding accuracy in 7/10 subjects using a cross-validation procedure applied in the predictive condition. Interestingly, similar accuracy (significant in 5/10 subjects) is obtained when the CSPs trained in the predictive condition are tested in the ambiguous condition. In agreement with this result, we find that the CSPs show very similar topographies in both conditions. These results shed a new light on the behavioral and EEG correlates of visuospatial attention in complex visual environments. This study demonstrates that alpha-power features could be used in brain computer interfaces to decode covert attention shifts in an environment containing ambiguous spatial information.

This work was published in *Frontiers in Human Neurosciences* [20].

6.8.2. Online extraction and single trial analysis of regions contributing to erroneous feedback detection

Participants: Eoin Thomas, Matthew Dyson, Laurence Casini, Boris Burle.

Understanding how the brain processes errors is an essential and active field of neuroscience. Real time extraction and analysis of error signals provide an innovative method of assessing how individuals perceive ongoing interactions without recourse to overt behaviour. This area of research is critical in modern Brain–Computer Interface (BCI) design, but may also open fruitful perspectives in cognitive neuroscience research. In this context, we sought to determine whether we can extract discriminatory error-related activity in the source space, online, and on a trial by trial basis from electroencephalography data recorded during motor imagery. Using a data driven approach, based on interpretable inverse solution algorithms, we assessed the extent to which automatically extracted error-related activity was physiologically and functionally interpretable according to performance monitoring literature. The applicability of inverse solution based methods for automatically extracting error signals, in the presence of noise generated by motor imagery, was validated by simulation. Representative regions of interest, outlining the primary generators contributing to classification, were found to correspond closely to networks involved in error detection and performance monitoring. We observed discriminative activity in non-frontal areas, demonstrating that areas outside of the medial frontal cortex can contribute to the classification of error feedback activity.

This work was published in *NeuroImage* [13].

BEAGLE Project-Team

7. New Results

7.1. Subspace Clustering Using Evolvable Genome Structure

We have developed an evolutionary algorithm to tackle the subspace clustering problem. Subspace clustering is recognized as more difficult than standard clustering since it requires to identify not only the clusters but also the various subspaces where the clusters hold. We propose to tackle this problem with a bio-inspired algorithm that includes many bio-like features like variable genome length and organization, functional and non-functional elements, and variation operators including chromosomal rearrangements. These features give the algorithm a large degree of freedom to achieve subspace clustering with satisfying results on a reference benchmark with respect to state of the art methods. One of the main advantages of the approach is that it needs only one subspace clustering ad-hoc parameter: the maximal number of clusters. This is a single and intuitive parameter that sets the maximal level of details of the clustering, while other algorithms require more complicated parameter space exploration. The other parameters of the algorithm are related to the evolution strategy (population size, mutation rate, ...) and for them we use a single setting that turns out to be effective on all the datasets of the benchmark.

This work has been presented at the main conference for genetic & evolutionary computation, GECCO [31], where it received the best paper award and during the EvoEvo Workshop of ECAL 2015 [35].

7.2. Epigenetic inheritance speeds up evolution of artificial organisms

DNA is not the sole medium by which parents transmit information to their offspring. Epigenetic inheritance, in particular, is based on the partial transmission of the cellular state of the parental cell to its descendants. Although the reality of epigenetic inheritance is now firmly established, whether it has an influence on the long term evolutionary process is still subject to debate. To address this question, we used the RAevol extension of the Aevol simulator developed in the team, and defined 4 scenarios with static or dynamic environments and with or without epigenetic inheritance. Simulations in dynamic environments show that protein inheritance indeed increases the rate of evolution on the long term. But they also show that it impedes evolution in its very first stages. This negative effect can be explained by instabilities generated by the interference between the two inheritance mediums. On the opposite, the long term gain can be explained by protein inheritance reducing the constraints on the genetic regulation network.

This work has been published in the article [33].

7.3. In silico evolution improves statistical models of genome dynamics

Using Aevol, we have proved that statistical frameworks published in the last twenty years for inferring evolutionary genome rearrangements are flawed in two ways. First, they mistranslated a null hypothesis on a uniform breakage model, and second, they assumed that genomic breakable regions are known *a priori*. We propose ways to correct these flaws by combining mathematical approaches, simulations, observations and validation on real genomic data. The results will be of interest for an audience from evolutionary biology, computational biology, bioinformatics and mathematics. We successively show that:

- a truly uniform hypothesis on rearrangement breakages leads to a model with an equilibrium intergene size distribution that fits the measured one on diverse genomes,
- estimations based on the flawed uniform breakage model completely fail on simulations with the truly uniform model,
- coherently with previous studies the flawed, and to a lesser extent, the truly uniform model are rejected on amniote genomes if breakable regions are identified with intergenic regions,
- co-estimating the number of breakable regions with the rearrangement distance gives coherent values on amniote genomes.

A paper reporting these results has been submitted by Priscila Biller, Carole Knibbe and Eric Tannier.

7.4. Temperature-induced variation in gene expression burst size in metazoan cells

Gene expression is an inherently stochastic process, owing to its dynamic molecular nature. Protein amount distributions, which can be acquired by cytometry using a reporter gene, can inform about the mechanisms of the underlying microscopic molecular system. By using different clones of chicken erythroid progenitor cells harboring different integration sites of a CMV-driven mCherry protein, we investigated the dynamical behavior of such distributions. We show that, on short term, clone distributions can be quickly regenerated from small population samples with a high accuracy. On longer term, on the contrary, we show variations manifested by correlated fluctuation in the Mean Fluorescence Intensity. In search for a possible cause of this correlation, we demonstrate that in response to small temperature variations cells are able to adjust their gene expression rate: a modest (2 °C) increase in external temperature induces a significant down regulation of mean expression values, with a reverse effect observed when the temperature is decreased. Using a two-state model of gene expression we further demonstrate that temperature acts by modifying the size of transcription bursts, while the burst frequency of the investigated promoter is less systematically affected. For the first time, we report that transcription burst size is a key parameter for gene expression that metazoan cells from homeotherm animals can modify in response to an external thermal stimulus.

This work has been published in the article [11].

7.5. Deciphering the signalling networks of synaptic plasticity

Synaptic plasticity, i.e. adaptive modifications of synaptic strength between two neurons depending on their activity, is a main substrate for learning and memory. Experimentally, synaptic plasticity is commonly assessed using prolonged electrical stimulations. Since learning can arise from few or even a single trial, synaptic strength is expected to adapt rapidly. However, whether synaptic plasticity occurs in response to limited event occurrences remains elusive. To address this question, we started a collaboration with Laurent Venance Lab (experimental neuroscience, College de France, Paris). Combining experimental and modelling approaches, we investigated whether a low number of stimulations can induce plasticity in a major synaptic learning rule, spike-timing-dependent plasticity (STDP). It is known that 100 stimulations induce bidirectional STDP, i.e. spike-timing-dependent potentiation (tLTP) and depression (tLTD) at most central synapses. In rodent striatum, we found that tLTD progressively disappears when the number of stimulations is decreased (below 50 pairings) whereas tLTP displays a biphasic profile: tLTP is observed for 75-100 stimulations, absent for 25-50 stimulations and re-emerges for 5-10 stimulations. This tLTP, induced by very few stimulations (5-10) depends on the endocannabinoid (eCB) system. The eCB system has recently emerged as a pivotal pathway for synaptic plasticity because of its widely characterized ability to depress synaptic transmission on short- and long-term scales. Our result therefore indicate that eCBs also mediate potentiation of the synapse. To understand how eCB signaling may support such bidirectionality, we combined electrophysiology experiments with mathematical modeling. Our model describes the temporal kinetics of the biochemical species involved in a first signaling pathway leading from NMDAR to calmodulin and CaMKII with that of a second, distinct one that assembles mGluR and cytosolic calcium to eCB production and the resulting activation of CB1R. This demonstrated that STDP outcome is controlled by eCB levels and dynamics: prolonged and moderate levels of eCB lead to eCB-mediated long-term depression (eCB- tLTD) while short and large eCB transients produce eCB-mediated long-term potentiation (eCB-tLTP). Therefore, just like neurotransmitters glutamate or GABA, eCB forms a bidirectional system to encode learning and memory.

For reasons of publication strategy, our first co-publication on the subject presents our major experimental results [16]. A second article, featuring both experimental and modelling results, explains how the underlying signalling network can support the observed bidirectionality and is under submission.

7.6. Anomalous diffusion as an age-structured renewal process

Continuous-time random walks (CTRW) are one of the main mechanisms that are recurrently evoked to explain the emergence of subdiffusion in cells. CTRW were introduced fifty years ago as a generalisation of random walks, where the residence time (the time between two consecutive jumps) is a random variable. If the expectation of the residence time is defined, for instance when it is dirac-distributed or decays exponentially fast, one recovers “normal” Brownian motion. However, when the residence time expectation diverges, the CTRW describes a subdiffusive behavior. The classical approach to CTRW yields a non-Markovian (mean-field) transport equation, which is a serious obstacle when one wants to couple subdiffusion with (bio)chemical reaction. We took an alternative approach to CTRW that maintains the Markovian property of the transport equation at the price of a supplementary independent variable. We associate each random walker with an age a , that is the time elapsed since its last jump and describe the subdiffusive CTRW using an age-structured partial differential equations with age renewal upon each walker jump. In the spatially-homogeneous (zero-dimensional) case, we follow the evolution in time of the age distribution. An approach inspired by relative entropy techniques allows us to obtain quantitative explicit rates for the convergence of the age distribution to a self-similar profile, which corresponds to convergence to a stationary profile for the rescaled variables. An important difficulty arises from the fact that the equation in self-similar variables is not autonomous and we do not have a specific analytical solution. Therefore, in order to quantify the latter convergence, we estimate attraction to a time-dependent “pseudo-equilibrium”, which in turn converges to the stationary profile. The corresponding article is currently in press [38].

7.7. IGF-I signalling in neural stem cells during neurogenesis and aging

Downregulation of insulin-like growth factor (IGF) pathways prolongs lifespan in various species, including mammals. Still, the cellular mechanisms by which IGF signaling controls the aging trajectory of individual organs are largely unknown. Z. Chaker, in M. Holzenberg Lab (Centre de Recherche Saint-Antoine, Paris), asked whether suppression of IGF-I receptor (IGF-1R) in adult stem cells preserves long-term cell replacement, and whether this may prevent age-related functional decline in a regenerating tissue. Using neurogenesis as a paradigm, we showed that conditional knockout of IGF-1R specifically in adult neural stem cells maintained youthful characteristics of olfactory bulb neurogenesis within an aging brain. This in turn resulted in neuro-anatomical changes that improved olfactory function. To help interpret these results, we developed a mathematical model of stem cell differentiation using ordinary differential equations with time-dependent growth, division and death rates (to account for aging) and optimizing at each time step the amount of IGF-1R to maximize an experimentally-derived tissue efficiency criterion. The model predicts that decreased stimulation of growth in adults is indeed optimal for tissue aging. Thus, inhibiting growth and longevity gene IGF-1R in adult stem cells induced a gain-of-function phenotype during aging, marked by optimized management of cell renewal, and enhanced olfactory sensory function.

This work has been published in the article [14].

7.8. A novel model for leptin resistance

Leptin is a major hormone that regulates food intake and appetite in most mammals. Leptin increase in the blood tends to decrease the food intake and leptin is produced in proportion with fat depot. Leptin is therefore a simple probe that feed backs energy reserve to the brain and maintains a constant weight. It is a central hormone for this balance because KO mice without the leptin gene are quickly extremely obese. Also obese people (and animal) tend to have high concentration of leptin suggesting that after a certain point the brain ignores the leptin signal. We developed a mathematical model that explores this resistance developed by neural cells to leptin. This model predicts leptin resistance if food intake is artificially increased and predict a pathway to obesity by such mechanism. This work has been published by H. Soula (Beagle) in collaboration with F. Crauste (Dracula) with co-supervised PhD student Marine Jacquier[19].

7.9. Without eye contact, birds are Markovian!

Any social birds rely on acoustic messages to organize their daily activity (such as parenting and food foraging). In many occasions, birds are within earshot but not in visual contact and therefore should rely only on acoustic channel for this communication. In collaboration with the University of Saint-Etienne, we developed automatic extraction scripts that can detect birds vocalizations in a protocol of meeting with decreasing distance and with or without visual contact modality. Our worked showed that without visual contact birds are more synchronized and their vocal dynamics cannot be distinguished from a two state Markov chain. This markov property vanishes as soon as visual contact is reestablished. This work has been published in the main ethology journal: *Animal Behaviour*[23].

BIGS Project-Team

7. New Results

7.1. Stochastic modeling

7.1.1. Tumor growth modeling

Participants: P. Vallois, S. Wantz-Mézières
 External collaborator: J-S. Giet (IECL, Université de Lorraine)

A cancer tumor can be represented for simplicity as an aggregate of cancer cells, each cell behaving according to the same discrete model and independently of the others. Therefore to measure its size evolution, it seems natural to use tools coming from dynamics of population, for instance the logistic model. This deterministic framework is well-known but the stochastic one is worthy of interest. We work with a model in which we suppose that the size V_t at time t of the tumor is a diffusion process of the type :

$$\begin{cases} dV_t = r V_t \left(1 - \frac{V_t}{\kappa}\right) - c V_t + \beta V_t dB_t \\ V_0 = v > 0 \end{cases} \quad (9)$$

where $(B_t)_{t \geq 0}$ is a standard brownian motion starting from zero. Then (i) We define a family of time continuous Markov chains which models the evolution of the rate of malignant cells and approximate (under some conditions) the diffusion process (V_t) . (ii) We study in depth the solution to equation (1). This diffusion process lives in a domain delimited by two boundaries: 0 and $\kappa > 0$. In this stochastic setting, the role of κ is not so clear and we contribute to understand it. We describe the asymptotic behavior of the diffusion according to the values of the parameters. The tools we resort to are boundary classification criteria and Laplace transform of the hitting time to biological worthwhile level. We are able in particular to express the mean of the hitting time. We have an accepted paper in the journal Theory of Stochastic Processes [70].

7.1.2. A Multitype Branching Process Model of Heterogeneous Damages in vitro Cancer Cell Populations Treated by Radiotherapy

Participants: T. Bastogne, P. Vallois
 External collaborator: S. Pinel (CRAN, Université de Lorraine)

Cancer is the result of inter-dependent multi-scale phenomena and this is mainly why the understanding of its spread is still an unsolved problem. In integrative biology, mathematical models play a central role; they help biologists and clinicians to answer complex questions through numerical simulations and statistical analyses. The main issue here is to better understand and describe the role of cell damage heterogeneity and associated mutant cell phenotypes in the therapeutic responses of cancer cell populations submitted to a radiotherapy sessions during *in vitro* experiments. The cell heterogeneity is often described as randomness in mathematical modeling and different representations, such as Markov chains, branching processes and even stochastic differential equations, have been recently used. Conversely to these previous studies, which only focused on the steady-state responses of cell populations, we are interested by modeling the transient behavior after treatment and to identify the role of mutation heterogeneity in the global dynamic response of the cell populations. We propose to describe the survival response of an *in vitro* cancer cell culture treated by radiotherapy as a superposition of independent dynamics. Each cell is represented by a finite collection of cell mutation states with possible transitions between them. The population dynamics is given by an age-dependent multi-type branching process. From this representation, we obtain equations satisfied by the average size of the global survival population as well as the one of subpopulations associated with 10 mutation phenotypes. This work was presented via a poster communication in a international congress [40].

7.1.3. Modeling of response to chemotherapy in gliomas

Participant: S. Wantz-Mézières

External collaborators: M. Ben Abdallah, Yann Gaudeau, J.-M. Moureaux (CRAN, Université de Lorraine) and M. Blonski, L. Taillandier (CHU Nancy)

In the framework of a collaboration with neurologists (Luc Taillandier, Marie Blonski, CHU Nancy) and automaticians (Jean-Marie Moureaux, Yann Gaudeau, CRAN), around the thesis supervision of M. Ben Abdallah, our aim is to work out personalized therapeutic strategy in the monitoring of diffuse low-grade glioma patients. Regular monitoring with MRI are used to estimate the tumour volume ; we proposed a method by manual segmentation and statistically assessed its reproducibility by a subjective test. In order to design a decision-aid tool for the response to chemotherapy, our approach is phenomenological and we used simple regression tools to model and predict the cinetics of the tumour growth. We identified two different models. These results open up many perspectives, the main one being the modeling by multi-factor models, including biological and anatomopathological factors. This work is currently in progress.

7.1.4. Photodynamic therapy

Participant: C. Lacaux

External collaborators: T. Obara and M. Thomassin (CRAN, Université de Lorraine), L. Vinckenboch (Fribourg)

Our project focuses on an innovative application: the interstitial PDT for the treatment of high-grade brain tumors. This strategy requires the installation of optical fibers to deliver the light directly into the tumor tissue to be treated, while nanoparticles are used to carry the photosensitizer into the cancer cells. In order to optimize the intra-cerebral position of our optical fiber, two fundamental questions have to be answered: (1) What is the optimal shape and position of the light source in order to optimize the damage on malignant cells? (2) Is there a way to identify the physical parameters of the tissue which drive the light propagation?

Notice that we are obviously not the first ones to address these issues, and there is nowadays a consensus in favor of the algorithm proposed by L. Wang and S. L. Jacques for the simulation of light transport in biological tissues. However, our starting point is the observation that the usual methods slightly lack of formalism and miss formal representations that answer the questions of identifiability. In [16], in the framework of homogeneous biological tissues, we propose an alternative MC method to Wang's algorithm. Then we also propose a variance reduction method. Interestingly enough, our formulation also allows us to design quite easily a Markov chain Monte Carlo (MCMC) method based on Metropolis-Hastings algorithm and to handle the inverse problem (of crucial importance for practitioners), consisting in estimating the optical coefficients of the tissue according to a series of measurements. We have compared the proposed MC and MCMC method and Wang's algorithm: we see that our MC method is much more consistent. However, MCMC methods induce quick mutations, which paves the way to very promising algorithms in the inhomogeneous case. To handle the inverse problem, we derive a probabilistic representation of the variation of the fluence with respect to the absorption and scattering coefficients. This leads us to the implementation of a Levenberg-Marquardt type algorithm that gives an approximate solution to the inverse problem. Our results open the way for new improvements of Monte-Carlo methods in the context of light propagation. They should rather be seen as a starting point for new methods, including in inhomogeneous tissue. This work has been presented in several french seminars (Lille, Avignon, Paris Descartes, Orléans).

7.1.5. Time-changed extremal process as a random sup measure

Participant: C. Lacaux

External collaborator: G. Samorodnitsky (Cornell, USA)

In extreme value theory, one of the major topics is the study of the limiting behavior of the partial maxima of a stationary sequence. When this sequence is i.i.d., the unique limiting process is well-known and called the extremal process. Considering a long memory stable sequence, the limiting process is obtained as a simple power time change extremal process. Céline Lacaux and Gennady Samorodnitsky have proved in [38] that this limiting process can also be interpreted as a restriction of a self-affine random sup measure. In addition, they have established that this random measure arises as a limit of the partial maxima of the same long memory stable sequence, but in a different space. Their results open the way to propose new self-similar processes with stationary max-increments. Céline Lacaux has presented this work in an invited session of the international conference *Extreme Value Analysis* at Ann Arbor (June 2015).

7.1.6. Modulus of continuity of some conditionally sub-Gaussian fields, application to stable random fields

Participant: C. Lacaux

External collaborator: H. Biermé (Poitiers)

Hermine Biermé and Céline Lacaux maintain their collaboration on the study of anisotropic random fields. They have extended their previous work in the framework of conditionally sub-Gaussian random series. For such anisotropic fields, they have obtained a modulus of continuity and a rate of uniform convergence. Their framework enables the study of study e.g., Gaussian fields, stable random fields and multi-stable random fields. As invited speaker, Céline Lacaux has presented this work in the international conference *Adventure in Self-similarity* at Cornell University (June 2015) [17]. Another of their works in progress deals with the simulation of anisotropic Gaussian random fields and the estimation of their parameters using quadratic variations.

7.1.7. DNA sequences analysis

Participants: P. Vallois

External collaborators: A. Lagnoux and S. Mercier (Toulouse)

Here we want to determine the sequences that are biologically interesting and compare the results using the single local score H_n and using the pair $(H_n; L_n)$ where L_n is the length of the segment that realizes the best score. In that view, we work on the p-values associated to the observed samples.

7.1.8. Multicriteria Agregation for Health Economic Assessment

Participants: T. Bastogne, Y. Petot, P. Vallois

The framework of this work is the PhD thesis of Yann Petit. The first chapter of the thesis is a state of the art identifying the current challenges in medico-economic analyses. A review article should be submitted in spring 2016. We are currently working on the aggregation operators, based on fuzzy measures and the Choquet integral. Theoretical results have been obtained and a publication is planned to be submitted in the second half of 2016. Work continues by introducing probabilities. The next step will be to apply our theoretical results to real clinical cases.

7.1.9. Spatial and spatio-temporal modeling

Participant : A. Gégout-Petit

External collaborators: S. Li, L. Guerin-Dubrana (Inra Bordeaux)

In the framework of a collaboration with INRA Bordeaux about the esca-illness of vines, Anne Gégout-Petit with Shuxian Li developed different spatial models and spatio-temporal models for different purposes: (1) study the distribution and the dynamics of esca vines in order to tackle the aggregation and the potential spread of the illness (2) propose a spatio-temporal model in order to capture the dynamics of cases and measure the effects of environmental covariates. For this, we propose different hierarchic models with latent process associated with a bayesian inference. A part of the research has been submitted in a journal of biology [39]. Shuxian Li defended his PhD on December the 15th.

7.1.10. Stochastic modeling of fatigue crack propagation

Participants: R. Azaïs, A. Gégout-Petit

External collaborators: A.B. Abdesslem, M. Puiggali, M. Touzet (Bordeaux)

Fatigue crack propagation is a stochastic phenomenon due to the inherent uncertainties originating from material properties and environmental conditions. In a recent preprint [35], we propose to model and to predict the fatigue crack growth by a piecewise-deterministic Markov process associated with deterministic crack laws of the literature, namely the Paris-Erdogan equation defined by $da/dN = C(\Delta K)^m$ and the Forman equation given by $da/dN = C(\Delta K)^m / (K_c(1 - R) - \Delta K)$, where a is the crack length, N denotes the number of cyclic mechanical loads, ΔK is the range of the stress intensity factor and C , m , K_c and R are different parameters. We introduce a regime-switching model to express the transition between Paris' regime and rapid propagation which occurs just before failure. We also investigate the prediction of the fatigue crack path and its variability based on measurements taken at the beginning of the propagation. This work has also been presented in an international conference [25].

7.2. Estimation and control for stochastic processes

7.2.1. Inference for dynamical systems driven by Gaussian noises

Participant: S. Tindel

External collaborators: K. Chouk, A. Deya, Y. Hu, L. Khoa, D. Nualart, E. Nualart, F. Xu. (US)

The problem of estimating the coefficients of a general differential equation driven by a Gaussian process is still largely unsolved. To be more specific, the most general (\mathbb{R} -valued) equation handled up to now as far as parameter estimation is concerned is of the form:

$$X_t^\theta = a + \theta \int_0^t b(X_u) du + B_t,$$

where θ is the unknown parameter, b is a smooth enough coefficient and B is a one-dimensional fractional Brownian motion. In contrast with this simple situation, our applications of interest (motivated by some anomalous diffusion phenomenon in proteins fluctuations) require the analysis of the following \mathbb{R}^n -valued equation:

$$X_t^\theta = a + \int_0^t b(\theta; X_u) du + \int_0^t \sigma(\theta; X_u) dB_t, \quad (10)$$

where θ enters non linearly in the coefficient, where σ is a non-trivial diffusion term and B is a d -dimensional fractional Brownian motion. We have thus decided to tackle this important scientific challenge first.

To this aim, here are the steps we have focused on in 2015:

- Some limit theorems for general functionals of Gaussian sequences [6], or for functionals of a Brownian motion [3], which give some insight on the asymptotic behavior of systems like (2).
- Extension of pathwise stochastic integration to processes indexed by the plane in [1], which helps to the definition of noisy systems such as partial differential equations.
- Definition of new systems driven by a (spatial) fractional Brownian motion, such as the stochastic PDE considered in [37].
- The local asymptotic normality obtained for the system (2), which implies a lower bound on general estimators of the coefficient θ . This is the contents of the preprint [41].

7.2.2. Optimal estimation of the jump rate of a piecewise-deterministic Markov process

Participants: R. Azaïs, A. Muller-Gueudin

A piecewise-deterministic Markov process is a stochastic process whose behavior is governed by an ordinary differential equation punctuated by random jumps occurring at random times. In a recent preprint [33], we focus on the nonparametric estimation problem of the jump rate for such a stochastic model observed within a long time interval under an ergodicity condition. More precisely, we introduce an uncountable class (indexed by the deterministic flow) of recursive kernel estimates of the jump rate and we establish their strong pointwise consistency as well as their asymptotic normality. In addition, we propose to choose among this class the estimator with the minimal variance, which is unfortunately unknown and thus remains to be estimated. We also discuss the choice of the bandwidth parameters by cross-validation methods. This paper has also been presented in two national workshops.

7.2.3. Estimation and optimal control for the TCP process

Participant : R. Azaïs

External collaborators: N. Krell (Rennes), B. de Saporta (Montpellier)

In [33], we assume that the transition kernel is continuous with respect to the Lebesgue measure. This condition may be not satisfied in some applications, as for instance for the well-known TCP process that appears in the modeling of the famous Transmission Control Protocol used for data transmission over the Internet. As a consequence, we propose to investigate estimation followed by optimal control for this ergodic process. The particular framework defined by this process allows us to define an optimal policy for the estimation of its jump rate. We obtain at present an efficient method for estimating the moments of the conditional distribution of the inter-congestion times in an optimal way. This work is currently in progress.

7.2.4. Estimation of integrals from a Markov design

Participant : R. Azaïs

External participants: B. Delyon, F. Portier

Monte-Carlo methods for estimating an integral assume that the distribution of the random design is known. Unfortunately, some applications generate a design whose density function f is unknown. In this case, a solution is to perform the classical Monte-Carlo estimate of the integral by replacing f by a leave-one-out kernel estimator, and one may expect the convergence

$$\frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i)}{\hat{f}^{(-i)}(X_i)} \rightarrow \int \varphi d\lambda,$$

when the number n of independent data X_i goes to infinity. This difficult question has been investigated by François Portier and Bernard Delyon in a recent paper. We propose to extend this work to the more general case of a Markov design. This new model includes a large variety of applications, in particular in biology and climatology. Indeed, the data $(X_i, \varphi(X_i))$ are often obtained from a measuring instrument that is launched in its environment and thus follows a random walk in it. A paper on this work will be submitted soon.

7.2.5. Method of control for radiotherapy treatment using Decision Markov Processes

Participants : R. Azaïs, B. Scherrer, S. Tindel, S. Wantz-Mézières

In recent years, Bastogne, Keinj and Vallois designed a Markov model of the evolution of cells under a radiotherapy treatment. We are currently investigating the problem of optimizing the radiotherapy intensity sequence in order to kill as many cancerous cells as possible while preserving as many healthy cells, a problem that fits into the stochastic optimal control problem. Our preliminary efforts suggest that, since we are dealing with large populations of cells, the problem can be well approximated by a limit deterministic optimal control problem. We can solve this problem numerically with a Pontryagine approach, and symbolically (in the simplest cases) by identifying the critical points of some multivariate polynomials. The latter approach allows us to validate the fact that the former actually finds globally optimal solutions. This is a work in progress.

7.2.6. Numerical approximate schemes for large optimal control problems and zero-sum two player games

Participant: B. Scherrer

External collaborators: V. Gabillon, M. Ghavamzadeh, M. Geist, B. Lesner, J. Perolat, O. Pietquin, M. Tagorti

We have provided in [23] (ICML 2015) the first finite-sample analysis of the LSTD(λ) algorithm aimed at approximating the value of some fixed policy in a large MDP, through the approximation of the projected fixed point of the linear Bellman equation from samples. This analysis highlights the influence of the main parameter λ of the algorithm.

The long version of our previous work on the analysis of an approximate modified policy iteration for optimal control and its application to the Tetris domain is now published in JMLR [13]. The extension of this algorithm family for computing approximately-optimal non-stationary policies allows to improve the dependency with respect to the discount factor: we provide such improved bounds in [19], as well as examples that show that our analysis is tight (and cannot be further improved).

An original analysis of the variation of the approximate modified policy iteration for computing approximate Nash equilibria in the more general setting of two-player zero-sum games was published in ICML 2015 [22].

7.3. Algorithms and estimation for graph data

7.3.1. Modelling of networks of multiagent systems

Participant: A. Muller-Gueudin

External collaborators: A. Girard, S. Martin, I.C. Morarescu (CRAN, Nancy)

We relate here a starting of collaboration with researchers in Automatics in Nancy. We consider here networks, modeled as a graph with nodes and edges representing the agents and their interconnections, respectively. The objective is to study the evolution of the opinion of all the agents. The connectivity of the network, persistence of links and interactions reciprocity influence the convergence speed towards a consensus. The problem of consensus or synchronization is motivated by different applications as communication networks, power and transport grids, decentralized computing networks, and social or biological networks. We then consider networks of interconnected dynamical systems, called agents, that are partitioned into several clusters. Most of the agents can only update their state in a continuous way using only inner-cluster agent states. On top of this, few agents also have the peculiarity to rarely update their states in a discrete way by resetting it using states from agents outside their clusters. In social networks, the opinion of each individual evolves by taking into account the opinions of the members belonging to its community. Nevertheless, one or several individuals can change their opinions by interacting with individuals outside its community. These inter-cluster interactions can be seen as resets of the opinions. This leads us to a network dynamics that is expressed in term of reset systems. We suppose that the reset instants arrive stochastically following a Poisson renewal process. We have an accepted paper in the journal IEEE Transactions on Automatic Control [10].

7.3.2. Microbial interaction inference by network analysis

Participants: A. Gégout-Petit, A. Muller-Gueudin

External collaborators: A. Deveau (INRA Nancy), C. Raïssy (Inria Orpailleur)

The objective is to characterize microbial interactions in a particular environment: the truffles.

The truffle provides a habitat for complex bacterial communities. The role for bacteria in the development of truffles has been suggested but very little is known regarding the structure and the functional potential of the truffle's bacterial communities along truffle maturation. In a mathematical point of view, two micro-organisms are connected if they are not independent, conditionally to the other micro-organisms. Several models fit into this setting, especially the gaussian graphical models, the bayesians networks, and the graphical log-linear models. But the data, which can be zeros inflated, need developments and we have to proposed new models. Moreover, we are confronted to the problem that $n \ll p$, that is the sample size is much smaller that the number of variables ($n = 30, p = 200$). Last year, thanks to a financially supported project (PEPS), we have began a collaboration between statisticians and data-miners. The first approches have been notified in a report [31]. The statistical methodologies developed for this project could also be applied to human health (for instance identification of network between bacteria inside colon).

7.3.3. Lossy compression of unordered trees

Participant: R. Azaïs

External collaborators: J-B. Durand, C. Godin

A classical compression method for trees is to represent them by directed acyclic graphs. This approach exploits subtree repeats in the structure and is efficient only for trees with a high level of redundancy. The class of self-nested trees presents remarkable compression properties by this method because of the systematic repetition of subtrees. In particular, the compressed version of a self-nested tree T is a linear directed acyclic graph with only $1 + \text{height}(T)$ nodes. Unfortunately, it should be noted that trees without a high level of redundancy are often insufficiently compressed by this procedure. In a paper recently submitted for publication in an international conference [32], we introduce a lossy compression method that consists in computing in polynomial time for trees with bounded outdegree the reduction of a self-nested structure that closely approximates the initial data. We prove on a simulated dataset that the error rate of this lossy compression method is always better than the loss involved in a previous algorithm of the literature, while the compression rates are equivalent.

7.3.4. Inference for critical Galton-Watson trees from their Harris process

Participant: R. Azaïs

External collaborator: A. Genadot (Inria CQFD Bordeaux)

Galton-Watson trees are an elementary model for the genealogy of a branching population and thus play a central role in biology. Critical Galton-Watson trees are generated from a sibling distribution μ whose theoretical expectation $\sum k\mu(k)$ is equal to 1. Under this assumption, the well-known Harris process of a tree conditioned on having n nodes converges to a Brownian excursion characterized by the variance $\sigma^2 = \sum (k-1)^2\mu(k)$ of μ . We propose to exploit this asymptotic approximation to define a new estimate of the unknown parameter of interest σ^2 based on a least-square method. In particular, this new technique allows us to take into account the behavior of the Harris path with respect to its asymptotic theoretical expectation. In certain cases, we obtain a better confidence interval than the classical approach. A paper on this work is in preparation.

7.4. Regression and machine learning

7.4.1. Uniform asymptotic certainty bands for the conditional cumulative distribution function

Participants: S. Ferrigno, A. Muller-Gueudin

External collaborator: M. Maumy-Bertrand (IRMA, Strasbourg)

In this work with Myriam Maumy-Bertrand (IRMA, Strasbourg), we study the conditional cumulative distribution function and a nonparametric estimator associated to this function. The conditional cumulative distribution function has the advantages of completely characterizing the law of the random considered variable, allowing to obtain the regression function, the density function, the moments and the conditional quantile function. As a nonparametric estimator of this function, we focus on local polynomial techniques described in Fan and Gijbels [64]. In particular, we use the local linear estimation of the conditional cumulative distribution function.

The objective of this work is to establish uniform asymptotic certainty bands for the conditional cumulative distribution function. To this aim, we give exact rate of strong uniform consistency for the local linear estimator of this function. We show that limit laws of the logarithm are useful in the construction of uniform asymptotic certainty bands for the conditional distribution function. In particular, we use a single bootstrap to construct sharp uniform asymptotic bands of this estimator.

We illustrate our results with simulations and a study of fetal growth which is based on 694 fetuses (carefully selected by exclusion of multiple pregnancies, malformed, macerated or serious ill fetuses, or those with chromosomal abnormalities) autopsied in fetopathologic units of the "Service de foetopathologie et de placentologie" of the Maternité Régionale Universitaire (CHU Nancy, France) between 1996 and 2013.

We have presented our results in two international conferences with proceedings in Lille in June 2015 ("47èmes Journées de Statistique de la SFdS") [21] and London in December 2015 ("CM Statistics") [36].

7.4.2. Omnibus tests for regression models.

Participants: R. Azaïs, S. Ferrigno

External collaborator: M-J. Martinez Marcoux (LJK, Grenoble)

The aim of this collaboration with Marie-José Martinez Marcoux (LJK, Grenoble) is to compare, through simulations, several methods to test the validity of a regression model. These tests can be "directional" in that they are designed to detect departures from mainly one given assumption of the model (for example the regression function, the variance or the error) or global (for example the conditional distribution function). The establishment of such statistical tests require the use of nonparametric estimators various functions (regression, variance, cumulative distribution function). The idea would then be able to build a tool (package R) that allows a user to test the validity of the model it uses through different methods and varying parameters associated with modeling. This work is currently in progress.

7.4.3. Data analysis techniques: a tool for cumulative exposure assessment

Participant: J-M. Monnez

External collaborators : W. Kihal, B. Lalloué, C. Padilla, D, S. Zmirou-Navier

Everyone is subject to environmental exposures from various sources, with negative health impacts (air, water and soil contamination, noise, etc.) or with positive effects (e.g. green space). Studies considering such complex environmental settings in a global manner are rare. We propose to use statistical factor and cluster analyses to create a composite exposure index with a data-driven approach, in view to assess the environmental burden experienced by populations. We illustrate this approach in a large French metropolitan area. The study was carried out in the Great Lyon area (France, 1.2 M inhabitants) at the census Block Group (BG) scale. We used as environmental indicators ambient air NO₂ annual concentrations, noise levels and proximity to green spaces, to industrial plants, to polluted sites and to road traffic. They were synthesized using Multiple Factor Analysis (MFA), a data-driven technique without a priori modeling, followed by a Hierarchical Clustering to create BG classes. The first components of the MFA explained, respectively, 30, 14, 11 and 9% of the total variance. Clustering in five classes group: (1) a particular type of large BGs without population; (2) BGs of green residential areas, with less negative exposures than average; (3) BGs of residential areas near midtown; (4) BGs close to industries; and (5) midtown urban BGs, with higher negative exposures than average and less green spaces. Other numbers of classes were tested in order to assess a variety of clustering. We present an approach using statistical factor and cluster analyses techniques, which seem overlooked to assess cumulative

exposure in complex environmental settings. Although it cannot be applied directly for risk or health effect assessment, the resulting index can help to identify hot spots of cumulative exposure, to prioritize urban policies or to compare the environmental burden across study areas in an epidemiological framework [9].

7.4.4. Online Partial Principal Component Analysis of a Data Stream

Participant: J-M. Monnez

External collaborator: R. Bar (EDF, R & D)

Consider a data stream and suppose that each data vector is a realization of a random vector whose expectation varies with time, the law of the centered data vector being stationary. Consider the principal component analysis (PCA) of this centered vector called partial PCA. In this study are defined online estimators of direction vectors of the first principal axes by stochastic approximation processes using a data batch at each step or all the data until the current step. This extends a former result obtained by the second author by using one data vector at each step. This is applied to partial generalized canonical correlation analysis by defining a stochastic approximation process of the metric involved in this case using all the data until the current step. If the expectation of the data vector varies according to a linear model, a stochastic approximation process of the model parameters is used. All these processes can be performed in parallel.

Moreover, several incremental procedures of linear and logistic regression of a data stream were defined and tested and compared on existing batch data files and on simulated data streams.

7.4.5. Prognostic Value of Estimated Plasma Volume in Heart Failure.

Participant: J-M. Monnez

External collaborators: E. Albuissou, B. Pitt, P. Rossignol, F. Zannad (CHU, Nancy)

The purpose of this study was to assess the prognostic value of the estimation of plasma volume or of its variation beyond clinical examination in a post-hoc analysis of EPHEUS (Eplerenone Post-Acute Myocardial Infarction Heart Failure Efficacy and Survival Study).

Assessing congestion after discharge is challenging but of paramount importance to optimize patient management and to prevent hospital readmissions.

The present analysis was performed in a subset of 4,957 patients with available data (within a full dataset of 6,632 patients). The study endpoint was cardiovascular death or hospitalization for heart failure (HF) between months 1 and 3 after post-acute myocardial infarction HF. Estimated plasma volume variation (Δ ePVS) between baseline and month 1 was estimated by the Strauss formula, which includes hemoglobin and hematocrit ratios. Other potential predictors, including congestion surrogates, hemodynamic and renal variables, and medical history variables, were tested.

An instantaneous estimation of plasma volume at month 1 was defined and also tested.

Multivariate analysis was performed with stepwise logistic regression. Δ ePVS was selected in the model. The corresponding prognostic gain measured by integrated discrimination improvement was significant. Nevertheless, instantaneous estimation of plasma volume at month 1 was found to be a better predictor than Δ ePVS. LDA with mixed variables was also performed and confirmed these results.

In HF complicating myocardial infarction, congestion as assessed by the Strauss formula and an instantaneous derived measurement of plasma volume provided a predictive value of early cardiovascular events beyond routine clinical assessment. Prospective trials to assess congestion management guided by this simple tool to monitor plasma volume are warranted [4].

7.4.6. Death or hospitalization scoring for heart failure patients

Participant: J-M. Monnez

External collaborator: E. Albuissou (CHU Nancy)

The purpose of this study was to define an event - death or hospitalization - score for heart failure patients based on the observation of biological, clinical and medical historical variables. Some of them were transformed or winsorized. Two methods of statistical learning were performed, logistic regression and linear discriminant analysis, with a stepwise selection of variables. Aggregation of classifiers by bagging was used. Finally a score taking values between 0 and 100 was established.

7.4.7. A simultaneous stepwise covariate selection and clustering algorithm to discriminate a response variable with numerous values

Participant: J-M. Monnez

External collaborator: O. Collignon (LIH, Luxembourg)

In supervised learning the number of values of a response variable to predict can be high. Also clustering them in a few clusters can be useful to perform relevant supervised classification analyses. On the other hand selecting relevant covariates is a crucial step to build robust and efficient prediction models, especially when too many covariates are available in regard to the overall sample size. As a first attempt to solve these problems, we had already devised in a previous study an algorithm that simultaneously clusters the levels of a categorical response variable in a limited number of clusters and selects forward the best covariates by alternate minimization of Wilks' Lambda. In this paper we first extend the former version of the algorithm to a more general framework where Wilks's Lambda can be replaced by any model selection criterion. We also turned forward selection into stepwise selection in order to remove covariates while the procedure processes if necessary. Finally an application of our algorithm to real datasets from peanut allergy studies allowed confirming previously published results and suggesting new discoveries.

7.4.8. Statistical Analyses of Cell Impedance Signals in High-Throughput Cell Analysis

Participant: T. Bastogne, L. Batista

External Collaborator: El-Hadi Djermoune (Université de Lorraine, CRAN)

With the advent of high-throughput technologies, life scientists are starting to grapple with massive data sets, encountering challenges with handling, processing and moving information that were once the domain of astronomers and high-energy physicists [91]. We particularly focus the statistical analysis of large batch of time series with applications in the preclinical research in Cancerology. Our original contribution consists in developing new dynamical system identification methods suited to the processing of those type of data. System identification is a data-driven modeling approach more and more used in biology and biomedicine. In this application context, each assay is always repeated to estimate the response variability. The inference of the modeling conclusions to the whole population requires to account for the inter-individual variability within the modeling procedure. One solution consists in using mixed effects models but up to now no similar approach exists in the field of dynamical system identification. Therefore, our objective is to develop a new identification method integrating mixed effects within an ARX (Auto Regressive model with eXternal inputs) model structure. The parameter estimation step relies on the EM (Expectation-Maximisation) algorithm. First simulation results show the relevance of this solution compared with a classical procedure of system identification repeated for each subject. This work and derived was accepted in conference papers [34] [24] [18].

BIOCORE Project-Team

7. New Results

7.1. Mathematical methods and methodological approach to biology

7.1.1. Mathematical analysis of biological models

7.1.1.1. Mathematical study of semi-discrete models

Participants: Jean-Luc Gouzé, Frédéric Grogard, Ludovic Mailleret, Pierre Bernhard, Elsa Rousseau, Nicolas Bajeux.

Semi-discrete models have shown their relevance in the modeling of biological phenomena whose nature presents abrupt changes over the course of their evolution [96]. We used such models and analyzed their properties in several practical situations that are developed in Section 7.2.3, some of them requiring such a modeling to describe external perturbations of natural systems, and others to take seasonality into account. External perturbations of interacting populations occur when some individuals are introduced or removed from a natural system, which occurs frequently in pest control applications, either through the direct removal of pests, or through the introduction of biological control agents [71],[27]. Seasonality is an important property of most agricultural systems in temperate environments since the year is divided into a cropping season and a ‘winter’ season, where the crop is absent, as in our analysis of the sustainable management of crop resistance to pathogens [25] or in the dynamics of plant pathogens [50].

7.1.1.2. Model reduction and sensitivity analysis

Participants: Suzanne Touzeau, Jean-Luc Gouzé, Stefano Casagrande, Victor Bernal Arzola.

Analysis and reduction of biochemical models. Dynamic models representing complex biological systems with numerous interactions can reach high dimensions and include complex nonlinearities. A model reduction method based on process weighing and pruning was developed and implemented on various models (ERK signaling pathway, circadian rhythms in *Drosophila*) [41]. A global sensitivity analysis was performed to check the method robustness against parameter uncertainty and variability. This work is part of Stefano Casagrande’s ongoing PhD thesis and is also a collaboration with Bayer (Sophia-Antipolis).

Parameter identification in compartmental systems. In collaboration with F. Dayan (R&D Manager, Dassault Systèmes), we worked on practical problems of identifiability of parameters in linear pharmacokinetic models. This was the subject of the internship of V. A. Bernal [58].

7.1.2. Metabolic and genomic models

Participants: Jean-Luc Gouzé, Madalena Chaves, Ismail Belgacem, Olivier Bernard, Stefano Casagrande, Francis Mairet, Sofia Almeida.

7.1.2.1. Continuous models analysis

Piecewise quadratic systems for studying growth rate in bacteria. These new systems (first introduced in [82]) use an expression for growth rate that may depend on any number of variables and have several quadratic modes. Relative to the “classical” piecewise affine systems, this new formulation allows the existence of sliding motion as well as oscillatory behaviour for solutions at the thresholds where the vector fields are opposing [21].

Transcription and translation models in bacteria. We study detailed models of transcription and translation for genes in a bacterium, in particular the model of gene expression of RNA polymerase. We also study other models of the global cellular machinery. This is part of the PhD theses of Ismael Belgacem [11] and Stefano Casagrande, and done in collaboration with Inria IBIS project-team, in particular with D. Ropers.

Design of a bistable switch to control cellular uptake. In a joint work with Diego Oyarzún (Imperial College), we analyse the construction of a synthetic bistable system using an unbranched metabolic chain with a global enzyme regulator, as an application of [109]. Bistability can be achieved by choosing an appropriate pattern of regulation. Robustness conditions are given in terms of the promoter dynamic ranges to guarantee a bistable uptake flux [35].

A reduced model for the mammalian cell cycle. We focused on identifying and analyzing the main mechanisms behind the cell cycle and proposed a mathematical model to describe them. This reduced model successfully reproduces oscillatory behaviors including the progress towards a mitosis phase, and then mitosis itself, characterized by an increase in cyclin B. The model was the topic of a poster at the Signallife Workshop [68]. This is a collaboration with F. Delaunay (Ibiv Nice) in the framework of Labex Signallife.

7.1.2.2. Hybrid models analysis

Attractor computation using interconnected Boolean networks. During the visit of Daniel Figueiredo, we have worked on an extension of the method proposed in [83]. The idea is to not only use the attractors but also an appropriate set of strongly connected components in the computation of the asymptotic graph [115]. Numerical simulations show a great improvement in the problem of discarding spurious attractors.

Periodic orbits in non monotonic negative feedback circuits. We study the occurrence of periodic solutions in an n -dimensional class of negative feedback systems defined by smooth vector fields with a window of not necessarily monotonic activity. By circumscribing the smooth system by two piecewise linear ones, we show there exists an invariant toroidal region which contains a periodic orbit of the original smooth system [37].

7.1.2.3. Estimation and control

Optimal allocation of resources in a bacterium. We study by techniques of optimal control the optimal allocation between metabolism and gene expression during growth of bacteria [85], in collaboration with Inria IBIS project-team.

Control of a model of synthesis of a virulence factor. In collaboration with J.-A. Sepulchre (INLN Nice), we model the production of a virulence factor by a bacterium in a continuous stirred tank reactor. The production of this enzyme is genetically regulated, and degrades a polymeric external substrate into monomers. A nonlinear control is built [48].

7.2. Fields of applications

7.2.1. Bioenergy

7.2.1.1. Modelling microalgae production

Participants: Olivier Bernard, Antoine Sciandra, Frédéric Grogard, Ghjuvan Grimaud, Quentin Béchet, David Demory, Hubert Bonnefond, Jean-Philippe Steyer, Francis Mairet.

Experimental developments

Experiments have been carried out to study the effects of nitrogen limitation on the lipid production in microalgae [23] and support model development. These experiments have been carried out in the Lagrangian simulator, under constant or periodic light and temperature, varying the total amount of light dose in the day. The response in terms of storage carbon (triglycerides and carbohydrates) has been measured and correlated to the environment fluctuations.

Other experiments were carried out to reproduce the light signal percept by a cell in a raceway pond [84], derived from hydrodynamical studies [92]. An electronic platform was developed to reproduce this high frequency light signal. The experiments show that the microalgae adapt their pigments to the average light that they have received [23]. Experiments with coloured light demonstrated that the growth rate results from the absorbed light, whatever its wavelength.

A new methodology to measure cell viability has been set up. This approach is very promising to distinguish between net and gross growth rate [20]. It was used in the models to assess the impact of temperature on growth and mortality. The mortality turns out to increase exponentially with temperature. The effect of a short term temperature stress was also tested to understand the consequences of a temperature peak in a cultivation system. Finally, it was shown that microalgae can bear with temperature peaks above T_{\max} if they do not last too long [57].

On top of this, we set up a new experimental platform to carry out pilot experiments with solar light. The platform includes four raceways and the equipment to inoculate and harvest the microalgae [60]. We tested the impact of coloured film mimicking possible photovoltaic material. The collected data were used to calibrate models integrating the light spectrum [64].

These works have been carried out in collaboration with A. Talec, S. Rabouille, and E. Pruvost (CNRS/UPMC-Oceanographic Laboratory of Villefranche-sur-Mer LOV).

In collaboration with the IFREMER-PBA team (Nantes) we contributed to a study on the efficiency of dyes (BODIPY and Nile red) to quantify lipid content in microalgae [38].

Metabolism of carbon storage and lipid production

A macroscopic model for lipid production by oleaginous microalgae [7] has been previously proposed. This model describes the accumulation of neutral lipids (which can be turned into biofuel), carbohydrates and structural carbon. A review of the microalgal metabolism reconstruction [15] together with the associated metabolic models has been carried out [14]. A metabolic model has been set up and validated for the microalgae *Isochrysis luthera*. It predicts carbohydrate and lipid accumulation, under conditions of light/dark cycles and/or nitrogen deprivation [72], [1]. A model was developed to represent heterotrophic growth on a mixture of acetate and butyrate [39]. A metabolic model was set up, on the basis of the DRUM framework [1], in order to simulate autotrophic, heterotrophic and mixotrophic growth, and to determine how to reduce substrate inhibition.

Modelling the coupling between hydrodynamics and biology

In collaboration with the Inria ANGE team, a model coupling the hydrodynamics of the raceway (based on multilayer Saint Venant system) with microalgae growth was developed [79]. This model is supported by the work of ANGE aiming at improving the multi-layer Saint-Venant approach to more finely represent the hydrodynamics of the raceway [54].

Modelling the photosynthesis response to fast fluctuating light

The impact of hydrodynamics on the light perceived by a single cell was studied thanks to fluid dynamics simulations of a raceway pond [90]. The light signals that a cell experiences at the Lagrangian scale, depending on the fluid velocity, were then estimated. A Droop-Han model was used to assess the impact of light fluctuation on photosynthesis. A new model accounting for photoacclimation was also proposed [34]. Single cell trajectories were simulated, and the effect on photosynthesis efficiency was assessed using models of photosynthesis [91]. These results were compared to experimental measurements where the high frequency light was reproduced [84].

Modeling microalgae production processes

The integration of different models developed within BIOCORE [54], [19], [7] was performed to represent the dynamics of microalgae growth and lipid production in raceway systems, on the basis of the dynamical model developed to describe microalgal growth under light and nitrogen limitations. The strength of this model is that it takes into account the strong interactions between the biological phenomena (effects of light and nitrogen on growth, photoacclimation ...), temperature effect [78], [111] and the radiative transfer in the culture (light attenuation due to the microalgae).

Using these approaches, we have developed a model which predicts lipid production in raceway systems under varying light, nutrients and temperature [107]. This model is used to predict lipid production in the perspective of large scale biofuel production [54]. It was also used to assess the microalgal production potential in France,

when taking into account the actual meteorology on a 2.5 degree grid, for 2012, the use of lands, slope, proximity of nutrients and CO₂ [93].

In the framework of the ANR project Purple Sun, we developed a thermic model of a raceway pond within a greenhouse in order to estimate the culture temperature. We also included in the microalgae model the effect of light wavelength. This model has been calibrated on experimental data from LOV and has been used to support lighting strategy in order to optimize microalgal productivity (a patent on this process has been submitted).

Nitrogen fixation by diazotrophs

The fixation of nitrogen by *Croccosphaera watsonii* was represented with a macro metabolic model [87] quantifying the main fluxes of carbon and nitrogen in the cell. The model was calibrated and validated with the data of three experiments carried out with different duration of the light period and daily dose. Extension of the model were studied to include the effect of temperature [61].

This work is done in collaboration with Sophie Rabouille (CNRS-Oceanographic Laboratory of Villefranche-sur-Mer LOV).

Modelling thermal adaptation in microalgae

An extended statistical analysis was carried out on a database representing the temperature response of more than 200 microalgal species. First the model proposed by [78] turned out to properly reproduce the temperature response. A model was then extracted to predict the observed link between the cardinal temperatures. This led to the reduction of the parameter number down to 2, with still a good prediction capability.

We have used Adaptive Dynamics theory to understand how temperature drives evolution in microalgae. For a constant temperature, we have shown that the optimal temperature trait tends to equal the environment temperature. We then studied the case where temperature is periodically fluctuating [88]. We now use this method at the scale of the global ocean, validating our approach with experimental data sets from 194 species [42], [49].

7.2.1.2. Control and Optimization of microalgae production

On-line monitoring

Interval observers give an interval estimation of the state variables, provided that intervals for the unknown quantities (initial conditions, parameters, inputs) are known [86]. Several developments were carried out in this direction to improve the design and performances of interval observers, and accounting for a specific structure (*i.e.* triangular) or property (*i.e.* Input to State Stable), [104]. Interval observers were designed for the estimation of the microalgae growth and lipid production within a production process [101][54] and validated experimentally [100][29].

Optimization of the bioenergy production systems

Based on simple microalgae models, analytical optimization strategies were proposed. We assessed strategies for optimal operation in continuous mode using the detailed model for raceways [106], [107]. We first solved numerically an optimal control problem in which the input flow rate of the raceway is calculated such that the productivity in microalgae biomass is maximized on a finite time horizon. Then, we re-analysed the optimization problem and derived a simplified strategy to reach biomass productivities very near to the maximal productivities obtained with the optimal control. These approaches were extended to outdoor cultivation, considering a possible variable culture depth. The optimal strategy for both depth and dilution rate was proposed in order to better manage the process inertia and finally avoid over warming periods. This work was done during the doctoral stay of Riccardo de Luca (Univ. Padova).

We also propose a nonlinear adaptive controller for light-limited microalgae culture, which regulates the light absorption factor (defined by the ratio between the incident light and the light at the bottom of the reactor). We show by numerical simulation that this adaptive controller can be used to obtain near optimal productivity under day-night cycles [31].

Interactions between species

Large scale culture of microalgae for bioenergy involves a large biodiversity. Control of such systems requires to consider the interactions between the different species. Such systems involve bacteria and microalgae, and the competition between these organisms can have several equilibrium points, which can be studied with Monod, Contois and Droop type models [28].

In the framework of the ANR Facteur 4 project, we propose to drive this competition exploring different strategies in order to select species of interest.

We had formerly proposed an adaptive controller which regulates the light at the bottom of the reactor [102]. When applied for a culture with n species, the control law allows the selection of the strain with the maximum growth rate for a given range of light intensity. This is of particular interest for optimizing biomass production as species adapted to high light levels (with low photoinhibition) can be selected. We have also proposed a strategy based on light stresses in order to penalize the strains with a high pigment content and finally select microalgae with a low Chlorophyll content [12]. This characteristic is of particular interest for maximizing biomass production in dense culture. The strategy has been carried out at the LOV and eventually the productivity of *Tisochrysis lutea* was improved by 75% [62]. A patent on this strategy is under submission. Strategies to improve the temperature response have also been proposed. First we modelled the adaptive dynamics for a population submitted to a variable temperature [88]. This was used at the LOV to design experiments with periodic temperature stresses during 200 days aiming at enlarging the thermal niche of *Tisochrysis lutea*. It resulted in an increase by 2 degrees of the thermal niche [12].

Finally, in a more theoretical framework, we studied how to select as fast as possible a given species in a chemostat with two species at the initial instant. Using the Pontryagin maximum principle, we have shown that the optimal strategy is to maintain the substrate concentration to the value maximizing the difference between the growth rates of two species [73]. We now try to extend this result for n species with mutations.

7.2.2. Biological depollution

7.2.2.1. Control and optimization of bioprocesses for depollution

Participants: Olivier Bernard, Francis Mairet, Jean-Luc Gouzé.

We have considered the problem of global stabilization of an unstable bioreactor model (e.g. for anaerobic digestion), when the measurements are discrete and in finite number ("quantized"). These measurements define regions in the state space, wherein a constant dilution rate is applied. We show that this quantized control may lead to global stabilization: trajectories have to follow some transitions between the regions, until the final region where they converge toward the reference equilibrium [30].

Although bioprocesses involve an important biodiversity, the design of bioprocess control laws are generally based on single-species models. In [98], we have proposed to define and study the multispecies robustness of bioprocess control laws: given a control law designed for one species, what happens when two or more species are present? We have illustrated our approach with a control law which regulates substrate concentration using measurement of growth activity. Depending on the properties of the additional species, the control law can lead to the correct objective, but also to an undesired monospecies equilibrium point, coexistence, or even a failure point. Finally, we have shown that, for this case, the robustness can be improved by a saturation of the control.

Moreno [105] have proposed an optimal strategy for fed-batch bioreactor with substrate inhibition. Thanks to the Pontryagin maximum principle and the Hamilton-Jacobi equation, we have shown that the same strategy is still optimal when mortality is included in the model [75]. We have also studied the problem when the singular arc is non-necessary admissible everywhere (i.e. the singular control can take values outside the admissible control set). We have pointed out the existence of a frame point on the singular arc above which any singular trajectory is not globally optimal. Moreover, we have provided an explicit way for computing numerically the switching curves and the frame point [17].

7.2.2.2. Coupling microalgae to anaerobic digestion

Participants: Olivier Bernard, Antoine Sciandra, Jean-Philippe Steyer, Frédéric Grogard, Francis Mairet.

The coupling between a microalgal pond and an anaerobic digester is a promising alternative for sustainable energy production and wastewater treatment by transforming carbon dioxide into methane using light energy. The ANR Phycover project is aiming at evaluating the potential of this process [113], [112].

In a first stage, we developed models for anaerobic digestion of microalgae. Two approaches were used: first, a dynamic model has been developed trying to keep a low level of complexity so that it can be mathematically tractable for optimisation [97]. On the other hand, we have tested the ability of ADM1 [114] (a reference model which considers 19 biochemical reactions) to represent the same dataset. This model, after modification of the hydrolysis step [99] has then been used to evaluate process performances (methane yield, productivity...) and stability through numerical simulations.

We have proposed and analysed a three dimensional model which represent the coupling of a culture of microalgae limited by light and an anaerobic digester. We first prove the existence and attraction of periodic solutions. Applying Pontryagin's Maximum Principle, we have characterized optimal controls, including the computation of singular controls, in order to maximize methane production. Finally, we have determined numerically optimal trajectories by direct and indirect methods [74].

Finally, we have studied the coupling between three ecosystems: an anaerobic digester, a wastewater treatment pond (with microalgae and nitrifiers) and a microalgal pond. Different possible coupling configurations were tested in simulation. A numerical optimization was carried out to identify, depending on the choice of the objective function (energy production, pollution removal) the optimal arrangement between the three processes. The optimal volume for each process was then determined. This work has been carried out in the framework of the Phycover ANR project and was the subject of the internship of Ignacio Lopez (Universidad de Chile).

7.2.2.3. Life Cycle Assessment

Participants: Olivier Bernard, Jean-Philippe Steyer.

This work is the result of a collaboration with Arnaud Helias of INRA-LBE (Laboratory of Environmental Biotechnology, Narbonne) and Pierre Collet (IFPEN).

In the sequel of the pioneering life cycle assessment (LCA) work of [94], we continued to identify the obstacles and limitations which should receive specific research efforts to make microalgae production environmentally sustainable.

The improvements due to technological breakthrough (leading to higher productivities) have been compared to the source of electricity. It turns out that the overall environmental balance can much more easily be improved when renewable electricity is produced on the plant [36]. As a consequence, a new paradigm to transform solar energy (in the large) into transportation biofuel is proposed, including a simultaneous energy production stage. This motivated the design of the purple sun ANR-project where electricity is produced by semi transparent photovoltaic panels [77] under which microalgae are growing.

Finally, some work are aiming at normalising LCA for microalgae and proposing guidelines to make the LCA more easily comparable [22].

These works have been carried out in collaboration with E. Latrille and B. Sialve (INRA-LBE).

7.2.3. Design of ecologically friendly plant production systems

7.2.3.1. Controlling plant pests

Participants: Frédéric Grogard, Ludovic Mailleret, Suzanne Touzeau, Nicolas Bajoux.

Optimization of biological control agent introductions

The question of how many and how frequently natural enemies should be introduced into crops to most efficiently fight a pest species is an important issue of integrated pest management. The topic of optimization of natural enemies introductions has been investigated for several years [6] [108], unveiling the crucial influence of within-predator density dependent processes. Since parasitoids may be more prone to exhibit positive density dependent dynamics rather than negative ones, which are prevalent among predatory biocontrol agents, the current modeling effort consists in studying the impact of positive predator-predator interactions on the optimal introduction strategies (PhD of Nicolas Bajoux, [70], [71]).

The influence of the spatial structure of the environment on biological control efficacy has also been investigated; first results indicate that spatial structure tends to influence it in quite a same way as intra-specific competition does [27].

Connected research on the influence of space on the establishment of biological control agents is also being pursued both through computer simulations and laboratory experiments on parasitoids of the genus *Trichogramma*. This is the topic of the PhD thesis of Thibaut Morel Journel (UMR ISA) [13]; in particular, we showed how landscape connectivity or spatial heterogeneity shape establishment dynamics in spatially structured environments [33], [51], [40]. This research linked to invasion biology also led some of us to contribute with opinion or review contributions to a special issue on biological invasions, in connexion with the GdR Invabio [26], [32].

7.2.3.2. Controlling plant pathogens

Participants: Frédéric Grogard, Ludovic Mailleret, Suzanne Touzeau, Elsa Rousseau, Mélanie Bonneault.

Sustainable management of plant resistance

Because plants can get sick, we studied other plant protection methods dedicated to fight plant pathogens. One such method is the introduction of plant strains that are resistant to one pathogen. This often leads to the appearance of virulent pathogenic strains that are capable of infecting the resistant plants. It is therefore necessary to find ways to protect the durability of such resistances, which are a natural exhaustible resource. We looked at landscape scale spatial deployment strategies of resistant crops able to maximize crop yield [25], allowing for the modification of the spatial arrangement of resistant crops over cropping seasons, showing dramatic increases in crop yield in particular epidemic situations [25].

Experiments were also conducted in INRA Avignon, followed by high-throughput sequencing (HTS) to identify the dynamics of virus strains competing within host plants. Different plant genotypes were chosen for their contrasted effects on genetic drift and selection they induce on virus populations. Those two evolutionary forces can play a substantial role on the durability of plant resistance. Therefore we fitted a mechanistic-statistical model to these HTS data in order to disentangle the relative role of genetic drift and selection during within-host virus evolution [53], [69], [43], [44]. A stochastic model was also produced to simulate the effect of drift on the virus epidemiological dynamics and on the durability of qualitative resistances [59]. This is the topic of Elsa Rousseau's PhD thesis, and is done in collaboration with Frédéric Fabre (INRA Bordeaux) and Benoît Moury (INRA Avignon).

We also developed an epidemiological model describing the dynamics of root-knot nematodes in a protected vegetable cropping system, to design optimal management strategies of crop resistance [110]. The model was fitted to experimental and field data. Preliminary results show that alternating susceptible and resistant crops not only increased the resistance durability, but reduced the disease intensity over time [63].

Finally we developed an epidemiological model including non-conventional gene-for-gene interactions in crops, based on the phoma stem canker of oilseed rape, to assess the durability of crop resistance in the field and design efficient deployment strategies [65]. This ongoing work is part of the K-Masstec project, which also incorporates experimental and field studies in collaboration with BIOGER (INRA Grignon).

Eco-evolutionary dynamics of plant pathogens in seasonal environments

Understanding better pathogen evolution also requires to understand how closely related plant parasites may coexist. Such coexistence is widespread and is hardly explained through resource specialization. We showed that, in agricultural systems in temperate environments, the seasonal character of agrosystems is an important force promoting evolutionary diversification of plant pathogens [89]. The plant parasites reproduction mode may also strongly interact with seasonality. In this context, we investigated the special case of oak powdery mildew, an oak disease which is actually caused by a complex of two different species, combining original plant epidemic data with the semi-discrete seasonal plant epidemic model we introduced a few years ago [50] [95]

This work has been done in collaboration with Frédéric Hamelin (Agrocampus Ouest) during Anne Bisson's internship, Marie Laure Desprez Loustau and Frederic Fabre (INRA Bordeaux).

7.2.3.2.1. Optimality/games in population dynamics

Participants: Frédéric Grogard, Ludovic Mailleret, Pierre Bernhard.

Optimal foraging and residence times variations

A continued collaboration with Vincent Calcagno (UMR ISA) has yielded paper where we reanalyzed the so-called Marginal Value Theorem (MVT), first published in 1976 [80], [81]. Ongoing work aims at pointing how this latter theorem has been misused in some biological literature.

We also investigated the problem in foraging theory of evaluating the expected harvest of an animal when conspecifics may arrive on the same patch of resource in a stochastic fashion, specifically according to a Poisson process or a Bernoulli process. A joint article with Frédéric Hamelin (Agrocampus Ouest) has been submitted for publication.

With Marc Deschamps, similar questions were studied in theoretical economy in the context of a Cournot competition on a single market. Again, an article has been submitted for publication.

The handicap paradox

We have investigated the question of “how could evolution have reached a state characterized by the handicap paradox?” with the tools of adaptive dynamics. We have reached the conclusion that, if one accepts adaptive dynamics as a model of evolution, and our model of sexual selection, the handicap paradox equilibrium is indeed the limit state of evolution [18].

This work was conducted with Frédéric Hamelin (Agrocampus Ouest).

BONSAI Project-Team

7. New Results

7.1. Ancestral gene order reconstruction

In the field of **genomic rearrangement**, a topic of interest is to infer ancestral gene order from gene order known in extant species. The problem resumes to compute a set ancestral CARs (continuous ancestral regions) at a given node of a phylogeny. This work, initially published in a conference, was published this year in a journal [5].

7.2. Nonribosomal peptides

Norine is the unique and leading platform dedicated to computational biology analysis of nonribosomal peptides (NRPs). It is used by thousands of scientists all over the world to explore and better understand the diversity of the NRPs. To improve the data quality and quantity in Norine, we are now opening our resource to external contributors. To achieve this new challenge, we developed new tools (MyNorine, s2m) and communicate on our novelties.

- **Crowdsourcing.** To facilitate the submission of new nonribosomal peptides (NRPs) or modification of stored ones in Norine, we have developed a dedicated and user-friendly module named MyNorine [2]. It provides interactive forms to fill in the annotations with, for example, auto-completion and tools such as a monomeric structure editor. It has especially been designed for biologists and biochemists working on secondary metabolites to easily enrich the database with their own data.
- **Norine communication.** We advertise Norine by different promoting media. We organized an international workshop in Lille in October to teach biologists and biochemists how to annotate NRPs and their synthetases with bioinformatics tools such as Norine. It attracted 32 attendees from 8 countries. We participated, as invited contributors, to the special issue "Bioinformatics tools and approaches for synthetic biology" of the new journal Synthetic and Systems Biotechnology edited by KeAi Publishing, funded by Elsevier and Chinese Science Publishing & Media. Our article [6] describes the usefulness of Norine to discover novel nonribosomal peptides, with examples of biological results obtained thanks to Norine tools. More than 20 NRPs have already been submitted since September, proving the efficiency of our communication and usefulness and relevancy of Norine.
- **Monomeric structure.** The tool Smiles2Monomers (abbreviated s2m) infers efficiently and accurately the monomeric structure of a polymer from its chemical structure [1]. It is provided to the scientific community through the Norine website for on-line run or for download. Beside its utility to facilitate the annotation of new peptides, it allowed us to detect annotation errors in the Norine database.

7.3. High-throughput V(D)J repertoire analysis

High-throughput V(D)J repertoire analysis is an activity started in the group in 2012. As mentioned in previous reports, we produced a platform dedicated to analysing lymphocyte populations: Vidjil. Starting from DNA sequences, Vidjil is able to identify and quantify lymphocyte populations, visualise them and store metadata. Vidjil is now used routinely in Lille hospital and is also tested in other laboratories around the world.

With collaborators in Prague we used Vidjil in a retrospective study on patients suffering acute lymphoblastic leukemia [3]. The study identified a new measure of predicting relapse in patients, just a month after the diagnosis. This measure is simple as it relies on the diversity of the lymphocyte population.

7.4. Spaced seed coverage

In the field of spaced seed statistics these last two years, a new challenge is the selection of a set of spaced seeds that are at the same time sensitive, while providing a stable similarity measure for *alignment-free genomic sequence comparison*. One of the most stable estimators is the *coverage* provided by these seeds. We have proposed an efficient method to build the coverage automaton, in order to compute several statistics efficiently. This work was implemented in Iedera and published in the AISM journal [4].

7.5. Genome scaffolding with contaminated data

Scaffolding is a cornerstone in the assembly of genomes from next-generation sequencing data. It consists in ordering assembled sequences according to their putative order and orientation in the source genome. However, we are almost always in a setting where the genome is not known. Instead, order and orientation of sequences are inferred from partial information present in the sequencing data.

Unfortunately, sequencing data is noisy and often has contamination, i.e. a subset of the data which indicates a wrong genome order and/or orientation. We have investigated this effect and designed the first algorithm that explicitly models this contamination to better perform scaffolding.

This work appeared in the proceedings of the WABI 2015 conference [9] and has been accepted to the Bioinformatics journal, currently under revision. This work is in collaboration with K. Sahlin and L. Arvestad (KTH, Sweden).

7.6. Mining metatranscriptomic data

The team has recently developed the SortMeRNA software, which is a sequence analysis tool for filtering, mapping and OTU-picking NGS reads. The core algorithm is based on approximate seeds and allows for fast and sensitive analyses of nucleotide sequences. In [11], we demonstrate a computational technique for filtering ribosomal RNA from total RNA in metatranscriptomic data using it. Additionally, we propose a post-processing pipeline using the latest software tools to conduct further studies on the filtered data, including the reconstruction of mRNA transcripts for functional analyses and phylogenetic classification of a community using the ribosomal RNA. This work is a collaboration with Genoscope.

7.7. Structured RNAs

In many families of structured RNAs, the signature of the family cannot be characterized by a single consensus structure, and is mainly described by a set of alternate secondary structures. For example, certain classes of RNAs adopt at least two distinct stable folding states to carry out their function. This is the case of riboswitches, that undergo structural changes upon binding with other molecules, and recently some other RNA regulators were proven to show evolutionary evidence for alternative structure. The necessity to take into account multiple structures also arises when modeling an RNA family with some structural variation across species, or when it comes to work with a set of predicted suboptimal foldings. In this perspective, we have introduced the concept of RNA multistructures, that is a formal grammar based framework specifically designed to model a set of alternate RNA secondary structures. Continuing our work of 2014, we provide several motivating examples and propose an efficient algorithm to search for RNA multistructures within a genomic sequence. This work was published in [7].

CAPSID Project-Team

7. New Results

7.1. Annotating 3D Protein Domains

Many protein chains in the Protein Data Bank (PDB) are cross-referenced with EC numbers and Pfam domains. However, these annotations do not explicitly indicate any relation between EC numbers and Pfam domains. In order to address this limitation, we developed EC-DomainMiner, a recommender-based approach for associating EC (Enzyme Commission) numbers with Pfam domains [19]. EC-DomainMiner is able to infer automatically 20,179 associations between EC numbers and Pfam domains from existing EC-chain/Pfam-chain associations from the SIFTS database as well as EC-sequence/Pfam-sequence associations from UniProt databases.

7.2. Large-Scale Analysis of 3D Protein Interactions

As part of a continuing collaboration with a former doctoral student in the Orpailleur team, Anisah Ghoorah (now at the University of Mauritius), we used her KBDock database of all known PPIs to perform a large-scale statistical analysis of the secondary structure composition of known protein-protein binding sites [14]. This showed that some combinations of secondary structure features are significantly favoured, whereas other combinations are considerably dis-favoured. These findings could provide knowledge-based rules for the prediction of unsolved protein-protein interactions.

7.3. Predicting Drug Side Effects

Together with Harmonic Pharma SAS (a LORIA / Inria spin-out company), we developed the “GESSE” method for proposing new uses for existing therapeutic drug molecules by associating the Gaussian shapes of known drug molecules with their clinically observed side-effects [15].

7.4. Modeling a GPCR Receptor Complex

In collaboration with the BIOS team (INRA Tours) and the AMIB team (Inria Saclay – Île de France) we used our Hex protein docking software to help model a multi-component G-protein coupled receptor (GPCR) complex [12]. The resulting 3D structure was shown to be consistent with the known experimental data for the protein components of this trans-membrane molecular signaling system.

7.5. Modeling the Apelin Receptor

The Apelin receptor (ApelinR) is a GPCR which is important in regulating cardiovascular homeostasis. As part of an on-going collaboration with the Centre for Interdisciplinary Research (CIRB) at Collège de France, we modeled the interaction between the Apelin peptide and ApelinR [13]. This study provides new mechanistic insights which could lead to the development of therapeutic agents for the treatment of heart failure.

7.6. Identifying New Anti-Fungal Agents

In this collaboration with several Brazilian laboratories (at University of Mato Grosso State, University of Maringá, Embrapa, and University of Brasilia), we identified several novel small-molecule drug leads against the pathogenic fungus *Paracoccidioides lutzii* [17] which is a serious health threat, especially in Brazilian hospitals.

CARMEN Team

7. New Results

7.1. Inverse Problem

Electrocardiograms simulated by our group with a highly realistic and detailed forward model were used for several inverse-modeling studies [34], [33], [38], [35].

- **Stability analysis of the POD reduced order method for solving the bidomain model in cardiac electrophysiology:** In this work we show the numerical stability of the Proper Orthogonal Decomposition (POD) reduced order method used in cardiac electrophysiology applications. The difficulty of proving the stability comes from the fact that we are interested in the bidomain model, which is a system of degenerate parabolic equations coupled to a system of ODEs representing the cell membrane electrical activity. The proof of the stability of this method is based on an a priori estimate controlling the gap between the reduced order solution and the Galerkin finite element one. We present some numerical simulations confirming the theoretical results. We also combine the POD method with a time splitting scheme allowing a faster solution of the bidomain problem and show numerical results. Finally, we conduct numerical simulation in 2D illustrating the stability of the POD method in its sensitivity to the ionic model parameters. We also perform 3D simulation using a massively parallel code. We show the computational gain using the POD reduced order model. We also show that this method has a better scalability than the full finite element method.
- **In silico assessment of drug effects on human embryonic stem cells derived cardiomyocytes electrical activity:** Computational modeling and simulation is extensively used to investigate diseases in cardiac electrophysiological activity and also drug effects, side effects and interactions. Human embryonic stem cell-derived cardiomyocytes (hESC-CMs) have been recently considered as a promising tool in regenerative medicine: their major role in repairing damaged tissue is due to pluripotency and ability to differentiate. These pluripotent cells are also used in early stages of drugs development. Pharmaceutical companies use the MultiElectrode Array (MEA) device in order to perform many in vitro experiments on hESC-CMs. The goal of our study is to derive a mathematical model and to simulate these in vitro experiments.
- **Sensitivity of the Electrocardiography Inverse Solution to the Torso Conductivity Uncertainties:** Electrocardiography imaging (ECGI) is a new non invasive technology used for heart diagnosis. It allows to construct the electrical potential on the heart surface only from measurement on the body surface and some geometrical informations of the torso. The purpose of this work is twofold: First, we propose a new formulation to calculate the distribution of the electric potential on the heart, from measurements on the torso surface. Second, we study the influence of the errors and uncertainties on the conductivity parameters, on the ECGI solution. We use an optimal control formulation for the mathematical formulation of the problem with a stochastic diffusion equation as a constraint. The discretization is done using stochastic Galerkin method allowing to separate random and deterministic variables. The optimal control problem is solved using a conjugate gradient method where the gradient of the cost function is computed with an ad-joint technique. The efficiency of this approach to solve the inverse problem and the usability to quantify the effect of conductivity uncertainties in the torso are demonstrated through a number of numerical simulations on a 2D geometrical model. Our results show that adding $\pm 50\%$ uncertainties in the fat conductivity does not alter the inverse solution, whereas adding $\pm 50\%$ uncertainties in the lung conductivity affects the reconstructed heart potential by almost 50%.
- **Inverse Localization of Ischemia in a 3D Realistic Geometry: A Level Set Approach:** The reconstruction of cardiac ischemic regions from body surface potential measurements (BSPMs) is usually performed at a single time instant which corresponds to the plateau or resting phase of the cardiac

action potential. Using a different approach, we previously proposed a level set formulation that incorporates the knowledge of the cardiac excitation process in the inverse procedure, thus exploiting the spatio-temporal correlations contained in the BSPMs. In this study, we extend our inverse level-set formulation for the reconstruction of ischemic regions to 3D realistic geometries, and analyze its performance in different noisy scenarios. Our method is benchmarked against zero-order Tikhonov regularization. The inverse reconstruction of the ischemic region is evaluated using the correlation coefficient (CC), the sensitive error ratio (SN), and the specificity error ratio (SP). Our algorithm outperforms zero-order Tikhonov regularization, specially in highly noisy scenarios.

- Inverse problem in electrocardiography via the factorization method of boundary value problems: We present a new mathematical approach for solving the inverse problem in electrocardiography. This approach is based on the factorization of boundary value problems method. In this paper we derive the mathematical equations and test this method on synthetic data generated on realistic heart and torso geometries using the state-of-the-art bidomain model in the heart coupled to the Laplace equation in the torso. We measure the accuracy of the inverse solution using spatial Relative Error (RE) and Correlation Coefficient (CC).
- In the inverse problem in electrocardiology, the goal is to recover electrophysiological activity in the heart without measuring directly on its surface (without using catheter interventions). Note that today the inverse computation is frequently used by solving the quasi-static model. This model doesn't take into account the heart dynamic in time and may result in considerable errors in the reconstruction of the solution on the heart. In [1] we study a 3D numerical inverse problem constrained by the bidomain equations in electrocardiology. The state equations consisting in a coupled reaction-diffusion system modelling the propagation of the intracellular and extracellular electrical potentials, and ionic currents, are extended to further consider the effect of an external bathing medium. Thus, we demonstrate that the novel concept of applying electrophysiological data might be useful to improve noninvasive reconstruction of electrical heart activity. Finally, we present numerical experiments representing the effect of the heart dynamic on the inverse solutions. Moreover in [2], we study the stability result for the conductivities of the approximate bidomain model. The proof is based on the combination of a Carleman estimate obtained in [3] and certain weight energy estimates for parabolic systems.
- The static inverse ECG problem needs to solve the well known ill posed Cauchy problem for the Laplace equation. A new approach investigated in the team uses the method of factorization of boundary value problems. This method, studied for itself, provides in this context the computation of Dirichlet-Neumann operators as solution of a Riccati equation. Results have been presented at IEEE international symposium on biomedical imaging, New York april 16-19, 2015. Further investigations will be lead using more precise numerical methods to solve the Riccati equation. The non-linearity and time dependence of the coupling resistance between cardiac cells (gap junctions) is studied in the Liryc institute and thought to be of importance in the understanding of cardiac arrhythmias. The internship of Nhan Le Thanh was a first step to investigate their numerical simulation.

An Example of calculus on a torso is shown in figure 1

7.2. Cardiac Electromechanics

In [1] we study a coupled elliptic-parabolic system modeling the interaction between the propagation of electric potential and subsequent deformation of the cardiac tissue. The problem consists in a reaction-diffusion system governing the dynamics of ionic quantities, intra and extra-cellular potentials, and the linearized elasticity equations are adopted to describe the motion of an incompressible material. The coupling between muscle contraction, biochemical reactions and electric activity is introduced with a so-called active strain decomposition framework, where the material gradient of deformation is split into an active (electrophysiology-dependent) part and an elastic (passive) one. In this paper we prove existence of weak solutions to the underlying coupled reaction-diffusion system and uniqueness of regular solutions. We close with a numerical example illustrating the convergence of the method and some features of the model.

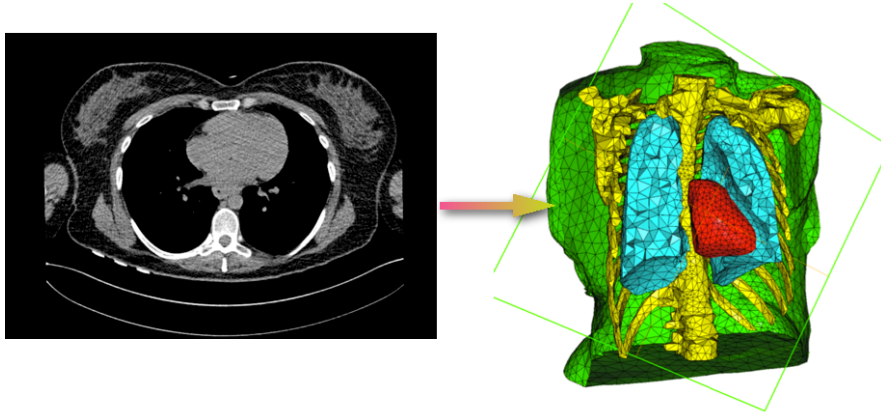


Figure 1. Mesh of the torso obtained with an IRM picture; this problem is link with the IDAM project

7.3. Cardiac Electrophysiology at the Microscopic Scale

We focused on establishing a microscopic model for cardiac electrophysiology simulations and proving the existence of a solution. We started with writing a mathematical proof allowing from well known physical equations and properties of the cardiac tissue to establish the model. Then, we worked on a variational formulation of the problem, and describing a weak solution of it. The idea is to compute energy estimates and to bound them so that we can extract a convergent sequence of functions in the appropriate Sobolev space. With my PhD advisor, we started to write an article about these two proofs. We also worked on CEPS code to implement some functionalities that will fit my requirements in a near future regarding the simulations we have to design. The main difficulty we identified is, provided we get a well defined geometry and mesh of cardiac cells, to implement the ionic flux between cells. First simulation of a simple "two-cells communication" problem will probably, if the results meet experimental observations, lead to another article. We also attended Imaged Based Biomedical Modelling 2015, a summer course organized by SCI institute (University of Utah), which was designed to give attendees guidelines about visualization and modelling, especially on cardio electrophysiology.

7.4. High order numerical scheme for ionic models

C. Douanla Ionti worked on time numerical schemes like Adams-Bashforth in order to have a high degree of convergence between an exact solution and the approximated solution. This method is a generalisation of Rush-Larsen scheme adapted for electrophysiology cardiac.

CASTOR Project-Team

5. New Results

5.1. Plasma boundary reconstruction

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras.

A new fast and stable algorithm has been developed for the reconstruction of the plasma boundary from discrete magnetic measurements taken at several locations surrounding the vacuum vessel. The resolution of this inverse problem takes two steps. In the first one we transform the set of measurements into Cauchy conditions on a fixed contour Γ_O close to the measurement points. This is done by least square fitting a truncated series of toroidal harmonic functions to the measurements. The second step consists in solving a Cauchy problem for the elliptic equation satisfied by the flux in the vacuum and for the overdetermined boundary conditions on Γ_O previously obtained with the help of toroidal harmonics. It is reformulated as an optimal control problem on a fixed annular domain of external boundary Γ_O and fictitious inner boundary Γ_I . A regularized Kohn-Vogelius cost function depending on the value of the flux on Γ_I and measuring the discrepancy between the solution to the equation satisfied by the flux obtained using Dirichlet conditions on Γ_O and the one obtained using Neumann conditions is minimized. The method presented here has led to the development of a software, called VacTH-KV, which enables plasma boundary reconstruction in any Tokamak (see [14]).

5.2. Free boundary - Transport Solver - Controller coupling

Participants: Cédric Boulbe, Blaise Faugeras, Jean François Artaud [IRFM CEA Cadarache], Vincent Basiuk [IRFM CEA Cadarache], Emiliano Fable [Max-Planck-Institut für Plasmaphysik, Garching], Philippe Huyn [IRFM CEA Cadarache], Eric Nardon [IRFM CEA Cadarache], Jakub Urban [IPP, Academy of Sciences of the Czech Republic, Prague].

Last year, a first version of the workflow coupling a free boundary equilibrium code, the European transport solver ETS and a plasma shape and position controller had been developed. In 2015, this new tool has been tested and improved. An experiment realized on the Tokamak TCV and called "yoyo" shot has been successfully simulated. This work has been realised in the framework of the Eurofusion Work Package: Code Development for integrated modelling project.

5.3. A finite element method with overlapping meshes for free-boundary toroidal plasma equilibria in realistic geometry

Participants: Holger Heumann, Francesca Rappetti.

Existing finite element implementations for the computation of free-boundary toroidal plasma equilibria approximate the flux function by piecewise polynomial, globally continuous functions. Recent numerical results for the self-consistent coupling of equilibrium and resistive diffusion in the spirit of Grad-Hogan suggest the necessity of higher regularity. Enforcing continuity of the gradient in finite elements methods on triangular meshes, leads to a drastic increase in the number of unknowns, since the degree of the polynomial approximation needs to be increased beyond four. Therefore existing implementations for the fixed boundary problem resort to (curvilinear) quadrilateral meshes and approximation spaces based on cubic Hermite splines. Fine substructures in the realistic geometry of a tokamak, such as air-gaps, passive structures and the vacuum vessel prevent the use of quadrilateral meshes for the whole computational domain, as it would be necessary for the free-boundary problem.

In this work we propose a finite element method that employs two meshes, one of quadrilaterals in the vacuum domain and one of triangles outside, which *overlap* in a narrow region around the vacuum domain. This approach gives the flexibility to achieve easily and at low cost higher order regularity for the approximation of the flux function in the domain covered by the plasma, while preserving accurate meshing of the geometric details exterior to the vacuum. The continuity of the numerical solution in the region of overlap is weakly enforced by relying on the mortar projection. A publication is in preparation.

5.4. Inverse transient plasma equilibrium problem

Participants: Holger Heumann, Jacques Blum.

The inverse transient plasma equilibrium problem aims at precomputing the trajectories of externally applied voltages in the poloidal field coils of a tokamak. A basic implementation of this problem in 2011/2012 in CEDRES++ during Holger Heumann's PostDoc at Inria, provided first insight into the capabilities and also difficulties of this approach. Application engineers are highly interested in this application, but realistic cases will require more evolved numerical methods to reduce the computational time and memory requirements. In 2014 we implemented the inverse transient plasma equilibrium problem in FEEQS.M to facilitate our search for better algorithms. In 2015 we started working on realistic test cases for the upcoming WEST tokamak. In order to make such problems accessible by the current version of our code, we had to split the time interval of interest into 5 subintervals, on which we solve 5 inverse problems. Only by the initial condition the problem on a subinterval is connected to its predecessor. Next we faced some serious convergence problem of the optimisation algorithms for some of these problems. These led us to do extensive benchmark runs with different optimisation algorithms and implementation, including both Gradient and SQP-type methods, either with handcoded or MATLAB-native implementations. As a result we envisage for 2016 the incorporation of the SQP implementation of Jean-Charles Gilbert, which seems to be perfectly adapted to optimal control problems such as ours. Another improvement was achieved in reducing the actual number of free control parameters and to replace piecewise linear control trajectories with high order polynomials.

5.5. High order for the axisymmetric magnetohydrodynamic equilibrium problem

Participants: Holger Heumann, Lukas Drescher [TU Berlin], Kersten Schmidt [TU Berlin].

We implemented a higher order finite element method (FEM) for solving numerically axisymmetric magnetohydrodynamic (MHD) equilibrium problems. The focus is on high accuracy and the capabilities of high-order FEM implementations for faster calculations. High order FEM for elliptic problems, such as the considered MHD equilibrium problem, is well established and understood. This work uses the hp-FEM software CONCEPTS developed at ETH Zürich/TU Berlin. Further, we developed a novel method for computing accurately the so-called flux surface averages, that are important in transient MHD calculations. This new method circumvents the expensive and very technical computation of line-integrals and fits seamlessly into the high order finite element method. A publication is in preparation.

5.6. Towards automated magnetic divertor design for optimal heat exhaust

Participants: Holger Heumann, Maarten Blommaert [FZ, Jülich (Germany)], Martine Baelmans [KU Leuven (Belgium)], Nicolas R. Gauger [TU, Kaiserslautern (Germany)], Detlev Reiter [FZ, Jülich (Germany)].

Avoiding excessive structure heat loads in future fusion tokamaks is regarded as one of the greatest design challenges. In this joint effort, we aim at developing a tool to study how the severe divertor heat loads can be mitigated by reconfiguring the magnetic confinement. For this purpose, the free boundary equilibrium code FEEQS.M was integrated with a plasma edge transport code to work in an automated fashion. A practical and efficient adjoint based sensitivity calculation was proposed to evaluate the sensitivities of the integrated code. The sensitivity calculation was applied to a realistic test case and compared with finite difference sensitivity calculations.

The integration of the free boundary equilibrium solver FEEQS.M allowed to assess the validity of a previous simplified model introduced by M. Bloomaert. It was found that the absence of plasma response currents significantly limits the accuracy of this simplified model.

The novel procedure was applied to obtain first results for the new WEST (Tungsten Environment in Steady-state Tokamak) divertor currently under construction in the Tore Supra tokamak at CEA. The sensitivities and the related divertor optimization paths are strongly affected by the extension of the magnetic model (see [24]).

5.7. Bohm boundary conditions

Participants: Richard Pasqueti, Sebastian Minjeaud.

Focusing on a minimal model proposed in the late 2000's by the IRFM (Cadarache), an algorithm has been proposed to enforce at the plates the inequality $M \geq 1$, where M is the parallel Mach number. The algorithm is implemented in the FBGKI code, but still requires improvements to enhance the robustness of the numerical method (see [18]).

5.8. High order approximation of dispersive equations and conservation of invariants

Participants: Richard Pasqueti, Sebastian Minjeaud.

Focusing on the Korteweg-de Vries (KdV) equation, algorithms have been proposed to handle high order derivative terms (third order for KdV) with C^0 elements and to preserve invariants (mass and momentum for KdV) through the time-scheme (see [33]).

5.9. Taylor-Galerkin stabilized Finite Element

Participants: José Costa, Boniface Nkonga.

The theoretical part of Taylor-Galerkin/Variational multi-scales (TG/VMS) strategy applied to MHD and reduced MHD modeling has been achieved last year. The final method amounts to adding in the finite element formulation, a self-adjoint operator associated to the most critical hyperbolic component of the system to be solved. The design of the critical contours and the identification of associated waves to be stabilized is problem dependent and related to the Jacobian matrix. This year we have continued the investigations for the design and improvement of the stabilization started in 2015. For application to plasma configurations with X-point, we have designed a numerical strategy that preserved the initial equilibrium without flows. The Bohm boundary condition on open flux walls has been formulated and is now under validation.

5.10. Toward full MHD numerical modeling with C^1 finite element.

Participants: José Costa, Giorgio Giorgiani, Hervé Guillard, Boniface Nkonga.

In this context the single fluid full MHD model is considered and the divergence free constraint on the magnetic field is achieved by introduction of a vector potential. The use of the vector potential has the additional advantage that the toroidal component is the magnetic flux of the Grad-Shafranov equilibrium. However, using the vector potential as variable introduces higher order derivatives in the system and classical C^0 finite elements cannot be directly applied. This is why our finite element strategies uses shape/test functions whose derivatives have global continuity in space (smooth finite elements). The global approach uses cross product shape/test functions between poloidal(2D) and the Toroidal(1D). In the 2D poloidal plane, discretization uses either quadrangular or triangular elements. In order to derive efficient strategies for the full MHD in the vector potential formulation, the Gauge condition on the vector potential and the boundary conditions have been enforced by penalization. For the Gauge condition it gives rise to element contributions but also boundary integrals that should be computed on curved surfaces that sometime fitted the magnetic surfaces. Equations are formulated in semi-conservative form such as to apply integration by parts. Therefore, boundary conditions can be viewed as evolution of fluxes or variables. Integral formulation on the boundary is very useful for higher order finite elements and also easier for the treatment of corners. Indeed in this context the boundary conditions are edge/surfaces oriented and boundary corners are driven by the neighborhood edge penalization. This strategy is the one that will be used for future developments.

2D Quadrangular Cubic Bezier Finite Elements:

This finite element has been used for a while for reduced MHD models in the software Jorek. Reduced MHD uses the projection of the momentum equation in a space orthogonal to the equilibrium. When full MHD models are used, the momentum equation needs to be projected in the equilibrium space and this projection should be consistent with the Grad-Shafranov equilibrium that is used to compute the initial state. This has been achieved by a proper computation of the $\mathbf{J} \times \mathbf{B}$ contribution in the momentum equation, taking into account the poloidal variation of the toroidal component of the magnetic field. After a detailed analysis, we have performed this year some implementations and numerical validations. An Inria report is under preparation.

2D Triangular Powell-Sabin Finite Elements:

In order to avoid some mesh singularities when using quadrangular meshes for complex geometries and flux surfaces shapes, triangular elements are a possible option. It is not so easy to derive smooth finite elements on triangles with reduced number of degrees of freedom (ddl). The Bell reduced-quintic finite elements we have considered in the previous years have too many unknowns (6 per vertex). Powell-Sabin splines are piecewise quadratic polynomials with a global C^1 -continuity and 3 unknowns per vertex, they have a local support, they form a convex partition of unity, they are stable, and they have a geometrically intuitive interpretation involving control triangles. Construction of the Powell-Sabin splines needs some geometrical tools that have been developed: Minimum area enclosing triangle of a set of control points (sequential and parallel). This construction is applied to each vertex of the triangular mesh and used to derive the local shape/test functions. These Powell-Sabin splines have been used successfully in the area of computer aided geometric design for the modeling and fitting of surfaces. We have used the Powell-Sabin (PS) splines for the approximation of elliptic partial differential equations (including Grad-Shafranov) in a rectangular domain. In this context, the optimal rate of convergence (order 3) has been recovered. This year, validations have been performed for hyperbolic 2D Euler equations with VMS stabilization. The context of the 3D toroidal geometries has been considered and implemented. Preliminary validations are satisfactory. An Inria report is also under preparation.

5.11. Genuinely multidimensional Riemann Solver

Participants: Jeaniffer Vides, Boniface Nkonga.

Multidimensional Riemann solvers were pioneered by Abgrall. Abgrall, Maire, Nkonga, Després and Loubere have extensively developed them especially as node-solvers for Lagrangian hydrodynamics. Another strain of work comes from explorations by Wendroff and Balsara who took a space-time approach. In this work, the resolved state is obtained via space-time integration over a wave model, just as was done by Wendroff and Balsara. However, an algebraic approach is used for the development of the fluxes. It is, therefore, shown that the multidimensional fluxes can be obtained by application of jump conditions at the boundaries of the wave model. The problem is of course over determined with the result that the shock jump conditions are only satisfied approximately in a least squares sense. Even so, this work gives us new perspective on multidimensional Riemann solvers. The literal satisfaction of the shock jump conditions (up to least squares approximation) makes it easier to understand multidimensional Riemann solvers as a natural extension of the one-dimensional Riemann solvers. Contributions have also been made on the development of a minimalist wave model, which might help in reducing dissipation. Further innovations are reported on the assembling of fluxes based on the structure of the wave model, and those innovations are potentially useful. For MHD the CT approach consists of constraining the transport of magnetic field so that the divergence is always kept zero. The method relies on exploiting the dualism between the flux components and the electric field. Since the electric field is needed at the edges of the mesh, the multidimensional Riemann solver can also provide the electric field. By running an extensive set of simulations, it is shown that the multidimensional Riemann solver is robust and can be used to obtain divergence-free formulations for MHD that perform well on several stringent calculations. The work performed this year was to improve this strategy by enriching of sub-structures the description of the strongly interaction of waves. These improvements were done in collaboration with the invited professor D. Balsara. This work has resulted in an article to be published in the Journal of Computational Physics in 2016.

5.12. Multi scales approximations of "Shallow water" flows.

Participants: Jeaniffer Vides, Boniface Nkonga, Sergey Gavriluk, Kseniya Ivanova.

The terminology "Shallow water" is used to characterize thin flows on curved surfaces. It is customary for this type of flows; to use the incompressible Navier-Stokes equations to asymptotically derive reduced models for the evolution of the depth integrated speed and the thickness of the flow. Reduced models are mainly hyperbolic and finite volume method are often used for their numerical approximation. Approximation strategies are generally structured as follow:

- Construction of a global coordinate system associated with an assumed analytical surface.
- Reduction of the model relatively to the global coordinate system
- Approximation of the surface by a finite number of elements.
- Approximation of the reduced model using the discrete surface.

In the context of real applications, it is presumptuous to expect an analytical formulation of the surface. From the data provided by observation satellites, we can usually extract a discrete description of the surfaces that drives thin flow. Therefore, it is more practical to use the discrete description as the starting point of the resolution strategy. This is the angle of approach that we have considered. We locally define two mesh scales: the element scale and the cell scale. The discrete mapping and the reduced model are defined at the element scale and the average values that evolve in time are defined at the cell scale. First applications have been successfully performed. Our efforts have been extended to include relevant physics at each scale, including sheared flows. We have used a multi-dimensional formulation. An Inria report is under preparation.

5.13. Asymptotic theory of reduced MHD models

Participant: Hervé Guillard.

In the study of fusion plasma, one of the fundamental model used for stability studies is the magnetohydrodynamic (MHD) model. Many theoretical and numerical works in this field use specific approximations of this model known as *reduced* MHD models. The derivation of these reduced MHD models has been formulated as a special instance of the theory of singular limit of hyperbolic system of partial differential equations with a large operator. This formulation allows to use the general results of this theory and to prove rigorously that reduced MHD models are valid approximations of the full MHD equations [29]. In particular, it is proven that the solutions of the full MHD system converge to the solutions of an appropriate reduced model. These results substantiate the intuitive physical idea that in the presence of a strong magnetic field, motion in the plane perpendicular to the plasma is nearly incompressible.

5.14. Finite volume approximations for fusion plasma

Participants: Hervé Guillard, Afeintou Sangam, Elise Estivals.

The MHD model used for plasma studies in tokamak is very often based on the magnetic vector potential form of the equations where the vector potential satisfies $\nabla \times \mathbf{A} = \mathbf{B}$ with \mathbf{B} the magnetic field and only a small number of numerical models use the conservative formulation based on \mathbf{B} . One of the shortcomings of this latter formulation is the necessity to enforce numerically the divergence free constraint on the magnetic field that can be difficult to achieve and/or computationally costly. Another difficulty is that the equilibrium solution of the MHD equation given by the Grad-Shafranov equation is not an exact solution of the discrete equation.

We have begun to investigate the use of the \mathbf{B} formulation for tokamak studies. The divergence free constraint is taken into account by a projection at each time step on a rotated gradient field. This step ensures a strict respect of the divergence free constraint while being extremely cheap since the scalar field is simply advected by the flow. Preliminary numerical experiments show that this approach can have some interest. The design of a well-balanced solver will be the next step of these studies.

CLIME Project-Team

6. New Results

6.1. Simulation, observation and state estimation for analysis and forecast

The objective of Clime is the merging of simulation and observations, with data assimilation methods, for state estimation in environmental applications. However, this aim previously requires, as seen in some of the next subsection, to collect the observations and carry out the simulations.

6.1.1. Assimilation of drifter data in the East Mediterranean Sea

Participants: Julien Brajard, Milad Fakhri [CNRS, Lebanon], Daniel Hayes [Oceanography Centre, Cyprus], Leila Issa [Lebanese American University, Lebanon], Laurent Mortier [LOCEAN], Pierre-Marie Poulain [Oceanography Institute of Trieste, Italy].

Surface velocity fields of the ocean in the Eastern Levantine Mediterranean are estimated by blending altimetry and surface drifters data. The method is based on a variational assimilation approach for which the velocity is corrected by matching real drifters positions with those predicted by a simple advection model, while taking into account the wind effect. The velocity correction is done in a time-continuous fashion by assimilating at once a whole trajectory of drifters using a sliding time window. A divergence-free regularization term is added to the cost function minimized during the assimilation process in order to constrain the velocity field. First results show that with few drifters, the method improves the estimation of the surface velocity: an eddy between the Lebanese coast and Cyprus is better assessed and the values of velocities along the Lebanese coast are more accurate.

6.1.2. Traffic simulation

Participants: Vivien Mallet, Vincent Aguiléra [CEREMA], Ruiwei Chen [CEREA].

The ANR project ESTIMAIR aims at propagating uncertainties in the complete simulation chain of air quality at urban scale. A key step in the chain lies in traffic assignment and the computation of the corresponding emissions. We take part to the simulation of traffic in the streets of Clermont-Ferrand metropolitan area, with the dynamic traffic assignment model LADTA. The simulations are evaluated against observations from loop counters and also against the simulations of the reference static model VISUM.

From the traffic assignment, the emissions are computed for nitrogen dioxide and particulate matter, using COPERT IV formulae. Preliminary work shows large uncertainties in the emissions due to the fleet composition.

6.1.3. Observation of noise pollution

Participants: Vivien Mallet, Raphaël Ventura, Valérie Issarny [MiMove], Pierre-Guillaume Raverdy [Ambientic], Fadwa Rebhi [MiMove].

Exposure to noise pollution is highly variable in space. As a consequence, it is very difficult to determine individual exposure using only numerical simulations of noise levels. Together with the MiMove Inria project-team, we take part to the SoundCity project that aims at collecting noise observations from smartphones and better evaluating the individual exposure. We assist MiMove in the development of an Android application that automatically senses noise along the day and collects the data (when the user agrees) for the improvement of simulated noise maps. Clime especially contributes to the calibration of the application. Comparisons between the measurements of smartphones and a sound meter allow us to estimate the bias of the main smartphones available on the market.

The SoundCity application was launched in July 2015 with Bernard Jomier, deputy mayor responsible for health, disability, and relations with Paris public hospital system, during a press conference organized by Paris City. The application received a positive coverage in the media, so that the application gained about 2500 users. About one million observations are collected every four days and ongoing work tries to process these data to correct Paris noise maps.

6.1.4. Evaluation of fire models

Participants: Jérémy Lefort, Vivien Mallet, Jean-Baptiste Filippi [CNRS].

In the field of forest fires risk management, important challenges exist in terms of people and goods preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires the evaluation of model performance.

We carry out the evaluation of several fire propagation models based on over 500 real fires. We use the data as they would be available in operational conditions, so as to avoid any tuning that would be incompatible with real-time forecasting. The study shows significant performance difference between the models, despite the poor data quality.

6.2. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, etc.) in order to improve the information provided by numerical models. However, the satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

6.2.1. Model error and motion estimation

Participants: Isabelle Herlin, Dominique Béréziat [UPMC].

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, an error term is included in the evolution equation of motion and a weak formulation of 4D-Var data assimilation is designed. The cost function to be minimized depends simultaneously on the initial motion field, at the beginning of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation. The approach is used to estimate geophysical forces (gravity, Coriolis, diffusion) from images in order to better assess the surface dynamics and forecast the displacement of structures like oilspill.

6.2.2. Tracking of structures from an image sequence

Participants: Isabelle Herlin, Yann Lepoittevin, Dominique Béréziat [UPMC].

The research concerns an approach to estimate velocity on an image sequence and simultaneously segment and track a given structure. It relies on the underlying dynamics' equations of the studied physical system. A data assimilation method is designed to solve evolution equations of image brightness, those of motion's dynamics. The method is for instance applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion, as seen on Fig. 2 .

Data assimilation is performed either with a 4D-Var variational approach or with a Kalman ensemble method. In the last case, the initial ensemble is obtained from a set of optical flow methods of the literature with various parameters values.

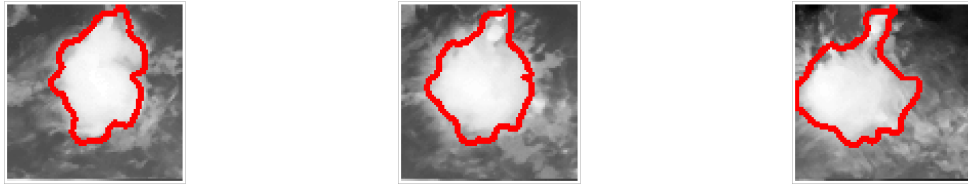


Figure 2. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

Various ways for representing the structures are studied and compared.

- For the variational approach, we consider: 1- a distance map modeling the tracked structures, which is added to the state vector, 2- anisotropic regularization terms, which are added to the cost function minimized during the assimilation process, 3- covariances between pixels, which are included in the background error covariance matrix.
- For the filtering approach, we focus either on domain decomposition or on explicit localization, which are both related to the displayed structures.

6.2.3. Applying POD on a model output database for defining a reduced motion model

Participants: Isabelle Herlin, Etienne Huot.

Dimension reduction may be obtained by determining a small size reduced basis computed by Proper Orthogonal Decomposition (POD) of a motion fields database and applying the Galerkin projection. This database is constructed for characterizing accurately the surface circulation of the studied area, so that linear combinations of the basis elements obtained by POD accurately describe the motion function observed on satellite image sequences. The database includes the geostrophic motion fields obtained from Sea Level Anomaly reanalysis maps that are available from the MyOcean European project website (<http://marine.copernicus.eu/>). Fig. 3 displays such SLA maps and the associated motion fields.

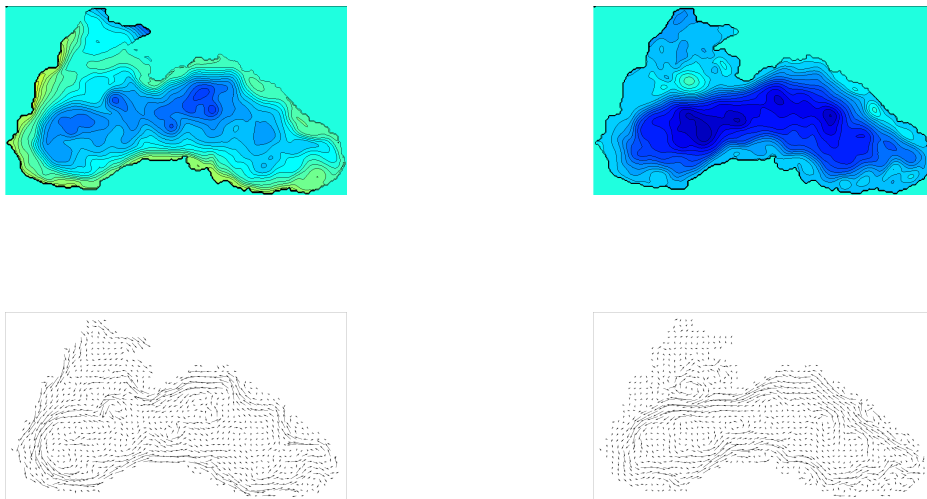


Figure 3. Top: reanalysis of SLA. Bottom: geostrophic motion.

Image assimilation with the POD reduced model allows estimating motion as displayed on Fig. 4 .

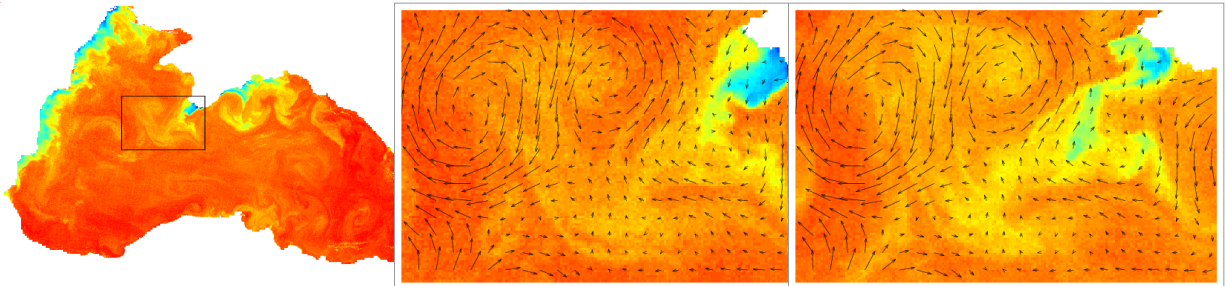


Figure 4. Zoom on a region of interest and motion estimation superposed on two consecutive images.

6.2.4. Rain nowcasting from radar image acquisitions

Participants: Isabelle Herlin, Yann Lepoittevin.

This research concerns the design of an operational method for rainfall nowcasting that aims at mitigating flash floods. The nowcasting method is composed of two main components:

- a data assimilation method, based on radar images, estimates the state of the atmosphere: this is the estimation phase.
- a forecast method uses this estimation to extrapolate the state of the atmosphere in the future: this is the forecast phase.

The method is transferred to the industrial company Weather Measures.

Current research concerns the use of object components in the state vector in order to get an improved motion estimation and a better localization of endangered regions. Assimilation of pluviometers measures in the nowcasting system is also investigated.

6.3. Uncertainty quantification and risk assessment

The uncertainty quantification of environmental models raises a number of problems due to:

- the dimension of the inputs, which can easily be 10^5 - 10^8 at every time step;
- the dimension of the state vector, which is usually 10^5 - 10^7 ;
- the high computational cost required when integrating the model in time.

While uncertainty quantification is a very active field in general, its implementation and development for geosciences requires specific approaches that are investigated in Clime. The project-team tries to determine the best strategies for the generation of ensembles of simulations. In particular, this requires addressing the generation of large multimodel ensembles and the issue of dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction is carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.3.1. Application of sequential aggregation to meteorology

Participants: Paul Baudin, Vivien Mallet, Gilles Stoltz [CNRS].

Nowadays, it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. When new observations are available, the meteorological centers also compute analyses. Therefore, we can apply the ensemble forecast of analyses, which consists in weighting the ensemble of forecasts to better forecast the forthcoming analyses. Before any forecast, the weights are updated with past observations and past forecasts. The performance of the aggregated forecast is guaranteed, in the long run, to perform at least as well as any linear combination of the forecasts with constant weights.

Ensembles of forecasts for mean sea level pressure, from the THORPEX Interactive Grand Global Ensemble, are aggregated with a forecast error decreased by 18% compared to the best individual forecast. The approach is also proved to be efficient for wind speed. The contribution of the ensembles (from different meteorological centers) to the performance increase are evaluated.

6.3.2. Sequential aggregation with uncertainty quantification and application to photovoltaics production

Participants: Paul Baudin, Vivien Mallet, Jean Thorey, Christophe Chaussin [EDF R&D], Gilles Stoltz [CNRS].

We study the aggregation of ensembles of solar radiations and photovoltaic productions. The aggregated forecasts show a 20% error decrease compared to the individual forecasts. They are also able to retrieve finer spatial patterns than the ones found in the individual forecasts (see Figure 5).

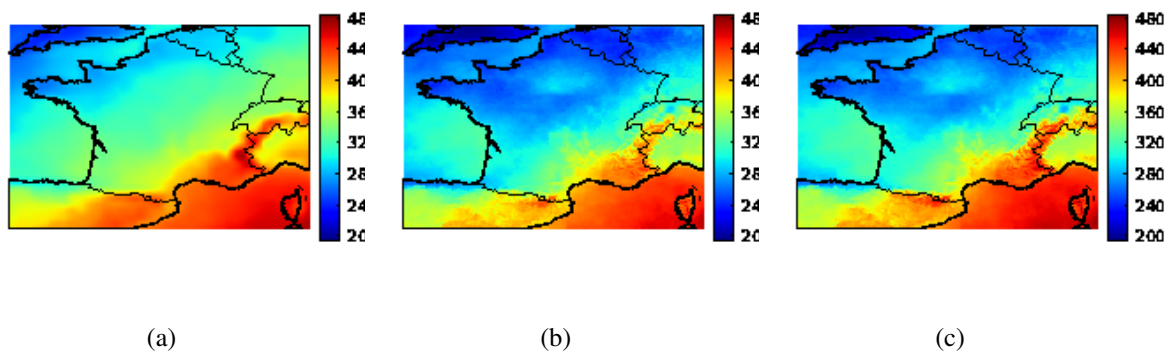


Figure 5. Yearly average of the map of downward shortwave solar radiation in Wm^{-2} , for an ensemble mean (a), for our aggregated forecasts (b) and observed (c).

An important issue is the estimation of the uncertainties associated with the aggregated forecasts. We devise a new approach to predict a probability density function or a cumulative distribution function instead of a single aggregated forecast. In practice, the aggregation procedure aims at forecasting the cumulative distribution function of the observations which is simply a Heaviside function centered at the observed value. Our forecast is the weighted empirical cumulative distribution function based on the ensemble of forecasts. The method guarantees that, in the long run, the forecast cumulative distribution function has a Continuous Ranked Probability Score (CRPS) at least as good as the best weighted empirical cumulative function with weights constant in time.

The CRPS is a classical score to evaluate the probabilistic forecasts. However, applying the CRPS on weighted empirical distribution functions (derived from the weighted ensemble) introduces a bias because of which minimizing the CRPS does not produce the optimal weights. Thus, we propose an unbiased version of the CRPS which relies on clusters of members and is strictly proper.

6.3.3. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard, Vivien Mallet, Irène Korsakissok [IRSN].

We carry out a sensitivity analysis of the dispersion of radionuclides during Fukushima disaster. We considered the dispersion at regional scale, with the Eulerian transport model Polair3D from Polyphemus. Simulations of the atmospheric dispersion of radionuclides involve large uncertainties originating from the limited knowledge of meteorological input data, composition, amount and timing of emissions and some model parameters. We studied the relative influence of each uncertain input on several outputs. In practice, we used the variance-based sensitivity analysis method of Sobol. This method requires a large number of model evaluations which are not achievable directly due to the high computational cost of the model. To circumvent this issue, we built a mathematical approximation of the model using Gaussian process emulation.

In previous studies, the uncertainties in the meteorological forecasts were crudely modeled by homogeneous and constant perturbations on the fields. Hence, we started investigating the use of ensembles of meteorological forecasts instead of just one base meteorological forecast. Including such ensembles allows to better represent the directions along which meteorological uncertainties should lie.

6.3.4. Fire risk assessment

Participants: Jérémy Lefort, Vivien Mallet, Jean-Baptiste Filippi [CNRS].

During days with extreme weather conditions, every wildland fire must be fought within minutes of its occurrence. This means that sufficient firefighting force is available at the right place and at the right time. In practice, firefighters wait at different critical locations, so that they can act quickly. For efficient preventive positioning of the firefighters, forecasting the risks of ignition of large fires is essential. This requires to predict where a fire may start, to estimate its potential size, to evaluate fighting scenarios and to anticipate which urban or protected areas may be under threat.

We designed a surrogate propagation model based on Gaussian process emulation of the model ForeFire. This surrogate model is fast enough to be run all over a region with high fire risk, e.g., Corsica. It can even be used for Monte Carlo simulations, with perturbations in the meteorological conditions and vegetation state, over Corsica. It is then possible to generate a risk map that identifies all the locations where a new fire can grow large.

6.3.5. Ensemble variational data assimilation

Participants: Julien Brajard, Isabelle Herlin, Marc Bocquet [CEREA], Jérôme Sirven [LOCEAN], Olivier Talagrand [LMD, ENS], Sylvie Thiria [LOCEAN].

The general objective of ensemble data assimilation is to produce an ensemble of analysis from observations and a numerical model which is representative of the uncertainty of the system. In a bayesian framework, the ensemble represents a sampling of the state vector probability distribution conditioned to the available knowledge of the system, denoted the a-posteriori probability distribution.

Ensemble variational data assimilation (EnsVar) consists in producing such an ensemble, by perturbing N times the observations according to their error law, and run a standard variational assimilation for each perturbation. An ensemble of N members is then produced. In the case of linear models, there is a theoretical guarantee that this ensemble is a sampling of the a-posteriori probability. But there is no theoretical result in the non-linear case.

Numerical experiments using non-linear numerical models suggest that the conclusion reached for linear models still stands for non-linear toy models.

The objective of this work is to study the ability of EnsVar to produce "good" ensemble (i.e. that sampled the a posteriori probability) on a more realistic model: a shallow-water model. Some statistical properties of the ensemble are presented, and the sensitivity to the main features of the assimilation system (number, distribution of observations, size of the assimilation window, ...) are also studied.

COFFEE Project-Team (section vide)

DEMAR Project-Team

5. New Results

5.1. Modelling and identification of the sensory-motor system

5.1.1. Implementation and Validation of a Stride Length Estimation Algorithm, Using a Single Basic Inertial Sensor on Healthy Subjects and Patients Suffering from Parkinson's Disease

Participants: Christine Azevedo Coste, Benoît Sijobert, Mourad Benoussaad [ENIT, Tarbes, France], Christian Geny [CHU Montpellier, Neurology, France], Jennifer Denys [stagiaire M2 STIC SANTE - DEMAR].

Providing a clinical oriented solution, our study presented a gyrometer and accelerometer based algorithm for stride length estimation. Compared to most of the numerous existing works where only an averaged stride length is computed from several IMU, or where the use of the magnetometer is incompatible with everyday use, our challenge here has been to extract each individual stride length in an easy-to-use algorithm requiring only one inertial sensor attached to the subject shank. Our results were validated on healthy subjects and patients suffering from Parkinson's disease (PD). Estimated stride lengths were compared to GAITRite walkway system data: the mean error over all the strides was less than 6 percents for healthy group and 10.3 percents for PD group. This method provides a reliable portable solution for monitoring the instantaneous stride length and opens the way to promising applications ([27]).

5.1.2. Dynamic mapping of upper limb tremor by muscle ultrasonography

Participants: Olivier Tassaert [stagiaire M1 - DEMAR / ICAR], Benjamin Gilles, Olivier Strauss [LIRMM], Christian Geny [CHU Montpellier, Neurology, France], Christine Azevedo Coste.

Focal treatment of action tremor by botulinum toxin injections has been inadequately investigated and at best provides modest relief with significant muscle weakness. Complexity of multi-joint tremulous movements results in non-individualized dosing regimens. Tremor is complex, especially in the upper extremity, and its manifestation can change depending on posture, task, and bodypart. Proper characterization of the tremor based on visual inspection alone is a daunting task for the clinician. Identification of the main trembling muscles task disturbing is challenging because many upper limb muscles are bi-functional. The performance of electromyographic (EMG) pattern-recognition based method in classifying movements strongly depends on arm positions and needs multiple measurements. High density-surface EMG (HD-sEMG) is a non-invasive promising technique to measure electrical muscle activity but has not been used for tremor research because deep muscles could not be investigated. Quantification of tremor dynamics by kinematics may be a feasible assessment and guidance tool which can be used to optimize injection conditions for focal tremor therapy. This approach is limited by the redundancy of the upper limb muscle organization. Contribution of synergistic muscles toward specific movements over multi joint systems may change with varying position of distal or proximal joints. The choice of injected muscles remains highly subjective and variable. In the study of Rahami, ten different arm or forearm muscles have been injected and improvement was mild and delayed and associated with muscle weakness. In recent years, muscle ultrasonography has become a promising tool for diagnosing neuromuscular disorders. This technique is a non-invasive, low-cost, imaging modality that may be used to characterize normal and pathological muscle tissue but also subtle muscular activity (fasciculations) in amyotrophic lateral sclerosis. The frequency of tremor remains stable during movement (3 to 8 Hz). We have initiated the investigation of the use of standard ultrasound as a technique to identify muscle groups responsible of upper limb tremor in patient with essential tremor or Parkinson's disease. The feasibility of the overall procedure has been validated: the acquisition procedure on patients, the possibility to track and segment the apparent motion in images using optical flow, and the ability to segment muscle groups by registering a 3D anatomical template.

5.1.3. *Understanding electrophysiological effects of direct electrical stimulation of the brain during wide awake surgery*

Participants: Marion Vincent, Olivier Rossel, Mitsuhiro Hayashibe, Hugues Duffau [CHU Montpellier], David Guiraud, François Bonnetblanc.

Direct electrical stimulation (DES) have been recently introduced in the neurosurgery of slow-growing and infiltrative brain tumors to guide the resection. By generating transient perturbations, this method allows the real-time identification of both cortical areas and subcortical networks that are essential for the function. Thus, as much as possible, non-functional tissue can be removed while minimizing the sequelae. However, there is much controversy as to whether the use of DES during wideawake surgery is the gold standard for studying the brain function. It is sometimes wrongly assumed that electrical microstimulation (EMS) and DES induce similar effects in the nervous tissues and have comparable behavioural consequences. Both of them are used to perform functional brain mapping: EMS for animal fundamental neuroscience experiments, and DES for neurosurgery patients. We tried to shed new light on electrical stimulation (ES) techniques in brain mapping by comparing EMS and DES [1]. In fact, their effects cannot be directly compared - especially in the electrophysiological domain. There is a gap between theory and practice for ES of the brain. We do not know exactly how ES and especially DES influence the electrophysiological state of networks in the brain; a strong biophysical rationale is lacking. In contrast, the gap between EMS and DES highlights the potential for new experimental paradigms in electrical stimulation for functional brain mapping. In view of this gap and recent technical developments in stimulator design, it may now be time to move towards alternative, innovative protocols. Moreover, the understanding of the electrophysiological effects of DES remains an open and key question. Intra-operative EEG (iEEG) recordings were studied to analyze if and how stimulation currents spread at distant sites. Data were collected during an awake brain surgery for one patient. We observed significant changes in the frequency content at different iEEG sites during DES [2]. Subcortical DES led to neuromodulation at more sites than cortical DES (Figure 4). This may be due to (i) a better conduction and propagation following the direct stimulation of large, myelinated axons and (ii) the greater current intensity in subcortical DES. Further research will have to characterize these aspects more carefully and apply cortical and subcortical DES with identical current intensities [30], [31].

5.1.4. *Functional Connectivity Analysis of Motor Imagery EEG signal for Brain-computer Interfacing Application: A preliminary study*

Participants: Saugat Bhattacharyya, Poulami Ghosh [Jadavpur University, India], Ankita Mazumder [Jadavpur University, India], D.n. Tibarewala [Jadavpur University, India], Mitsuhiro Hayashibe.

The human brain can be considered as a graphical network having different regions with specific functionality and it can be said that a virtual functional connectivity are present between these regions. These regions are regarded as nodes and the functional links are regarded as the edges between them. The intensity of these functional links depend on the activation of the lobes while performing a specific task(e.g. motor imagery tasks, cognitive tasks and likewise). The analysis of these networks are performed by using a very useful mathematical tool called graph theory. Graph theory basically represents the entire functional network with a number of nodes and edges between them and the amount of connectivity existing between two nodes is depicted by assigning weights to the edges between them. In this study we have tried to utilize functional connectivity between different parts of the human brain for classifying a motor imagery task.

Brain connectivity patterns can be determined by using two types of measures, namely, Bivariate and Multivariate. Here we have considered a multivariate measure known as multivariate autoregressive (MVAR) model. One of the most widely investigated connectivity measure is the Directed Transfer Function (DTF). This function basically computes the directional influences between any two given nodes. There are a number of theoretical indices for defining a graph. In this preliminary work, two indices, namely node strength and network density are measured from the DTF values. In the current study, the BCI competition Dataset III is used for computing different multivariate measures.

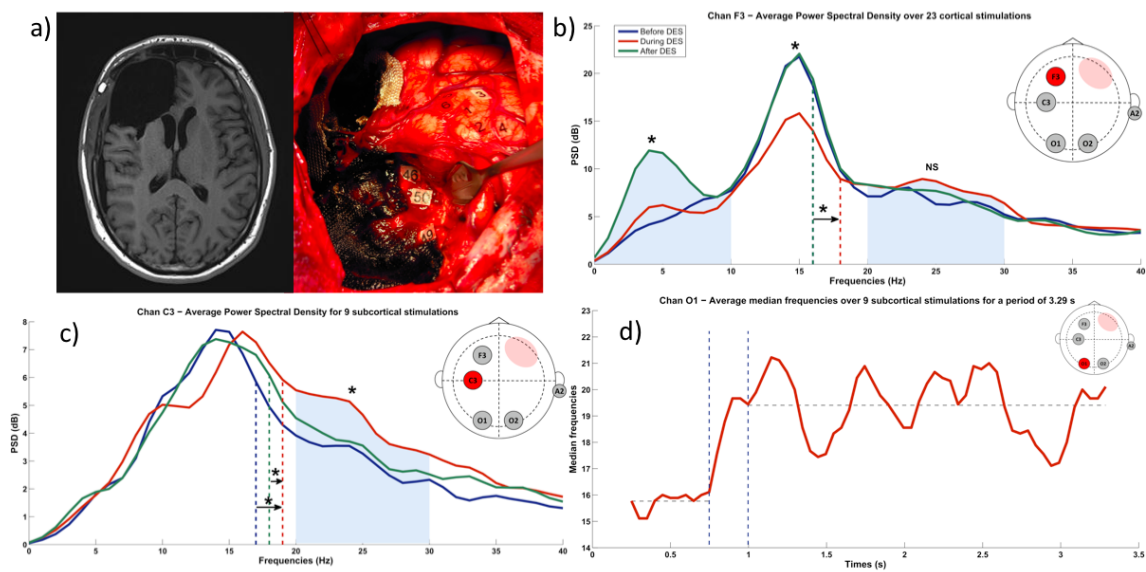


Figure 4. (a) Post-operative MRI of the patient's brain, showing the right frontal cavity and an intraoperative view of the brain with the main anatomical landmarks. (b) The mean PSD of the iEEG signal, on F3, before, during and after each period of cortical DES. (c) The mean PSD of the iEEG signal, on C3, before during and after each period of subcortical DES. (d) The moving window median frequency averaged over nine subcortical DESs periods for PSD measured at O1.

The inflow-outflow graph of subject 1 while imagining right hand movement in the first training set are given in Fig.5 . Fig. 5 (a) describes the amount of inflow of functional connectivity going out of all the 32 electrodes and these are color coded to indicate the intensity of these inflows. From Fig. 5 (a) it is quite evident that the inflows are maximum in the frontal, temporal and occipital lobes. Figure 5 (b) depicts the functional outflow from the nodes and in contrast to Fig.5 (a) it shows that the outflows are maximum from the Central lobe(Cz). In Fig 5 (c), the direction of the flow between different nodes are shown and it can be seen clearly that majority of the paths are going from Cz to different nodes of the frontal, parietal and temporal lobes.

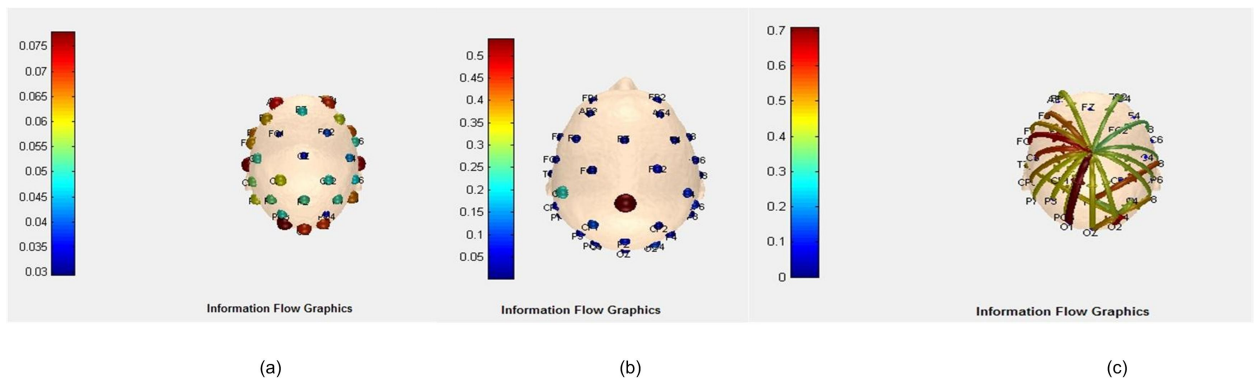


Figure 5. (a) Inflow graph, (b) Outflow graph and (c) Out to inflow graph of the functional connectivity network of the brain while imagining right hand movement.

5.1.5. A Generic Transferable EEG Decoder for Online Detection of Error Potential in Target Selection

Participants: Saugat Bhattacharyya, Amit Konar [Jadavpur University, India], D.n. Tibarewala [Jadavpur University, India], Mitsuhiro Hayashibe.

Detection of error from electroencephalography (EEG) signals as feedback while performing a discrete target selection task is beneficial for general Brain-computer Interfacing (BCI) systems including rehabilitative application. Error Related Potentials (ErrP) are EEG signals which occur when the participant observes an erroneous feedback from the system.

In this study, we have designed a novel scheme for detection of error feedback directly from the EEG signal. For this purpose, we have used a P300-speller dataset from the 'BCI Challenge @ NER 2015' competition hosted at Kaggle. The task involves the subject to select a letter of a word which is followed by a feedback period. The feedback period displays the letter selected and if the selection is wrong, the subject perceives it by the generation of ErrP signal. Our proposed system is designed to detect whether the feedback is erroneous or not. The decoder designed for this task is an ensemble of linear discriminant analysis, quadratic discriminant analysis and logistic regression classifier. The decoder is also transferable in nature as it should work with single-trial on new subject without any prior subject-specific training.

The block diagram of the BCI system adopted for online ErrP detection from input EEG signals is shown in Fig.6 . The system implements three main processes: i) Pre-processing of the signal, i.e., temporal filtering in the bandwidth [0.1, 10]Hz, ii) Extraction of relevant features corresponding to the mental state from the signal using savitzsky-golay filter and meta-data of the features, and iii) Classification of the features, using our proposed decoder, to detect the intention of the participant from two given states: *Error* and *No-Error*. A switch is incorporated in the design to detect the beginning of feedback period in the trials, which is marked

in the datasets. We have tested the online functionality of the BCI system on the test dataset provided in the website. To simulate a real-time condition, the EEG is continuously streamed until an onset of the feedback period is detected. On detection of the feedback period, the system extracts a pre-defined length of signal for further processing and the rest are rejected.

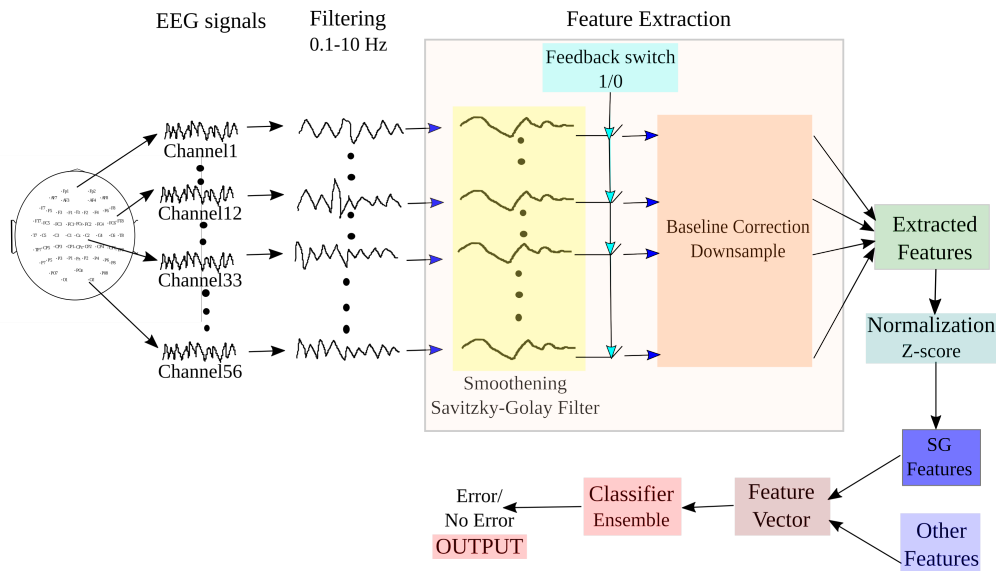


Figure 6. Block diagram of the BCI system adopted for online detection of Error Related Potentials from the input EEG

5.2. Synthesis and Control of Human Functions

5.2.1. FES-cycling and participation to Cyathlon competition

Participants: Christine Azevedo Coste, Benoît Sijobert, Charles Fattal [CRF DIVIO, Dijon, France], Antonio Padilha [UNB, Brasilia, Brazil], Emerson Fachin Martins [UNB, Brasilia, Brazil], David Andreu.

DEMAR and University of Brasilia will jointly participate with two SCI pilots to Cyathlon - FES-Bike competition. Cyathlon intends to promote assistive technologies during a competition. Two trikes will be adapted, an original control strategy will be proposed and two paraplegic individuals (one from Brazil and one from France) will be trained during the upcoming year. The protocol will be submitted to CPP ethical committee for agreement in the coming weeks (<http://freewheels.inria.fr/>).

5.2.2. PersoStim: A Personalized Closed-loop FES Control of Muscle Activation with Evoked EMG Feedback

Participants: Mitsuhiro Hayashibe, Zhan Li [University of Electronic Science and Technology of China], David Andreu, David Guiraud.

Functional electrical stimulation (FES) is a useful technique for restoring motor functions for spinal cord injured (SCI) patients. Muscle contractions can be artificially driven through delivery of electrical pulses to impaired muscles, and the electrical activity of contracted muscles under stimulus recorded by electromyography (EMG) is called M-wave. The FES-induced muscle activation which is represented by evoked EMG recordings can indicate the muscle state. Accurate control of muscle activation level by FES is the first step toward achieving more complicated FES control tasks.

A new FES closed-loop control strategy, EMG-feedback predictive control (EFPC), was developed to adaptively control stimulation pattern compensating to time-varying muscle state changes such as muscle fatigue and stimulation electrode detachment, along with the consideration of the personalized muscle responses to the electrical stimulation. This software manages a real-time FES system for control of muscle activation by online modulating pulse width of stimulus. The excitation muscle dynamics is modelled by Hammerstein system with stimulus pulse width and eEMG as input and output respectively. The model predictive control strategy is adopted to systematically produce the pulse width command of the stimulator. It is implemented together with Vivaltis portable stimulator. Four reference muscle activation patterns are provided to test and validate the real-time closed-loop FES control system. Real-time control results show promising control performances.

Recently, this software was demonstrated at the event of Rencontre Inria-Industrie 13/10/2015 at Bordeaux. <https://www.inria.fr/centre/bordeaux/innovation/rii-sante/demonstrations2>

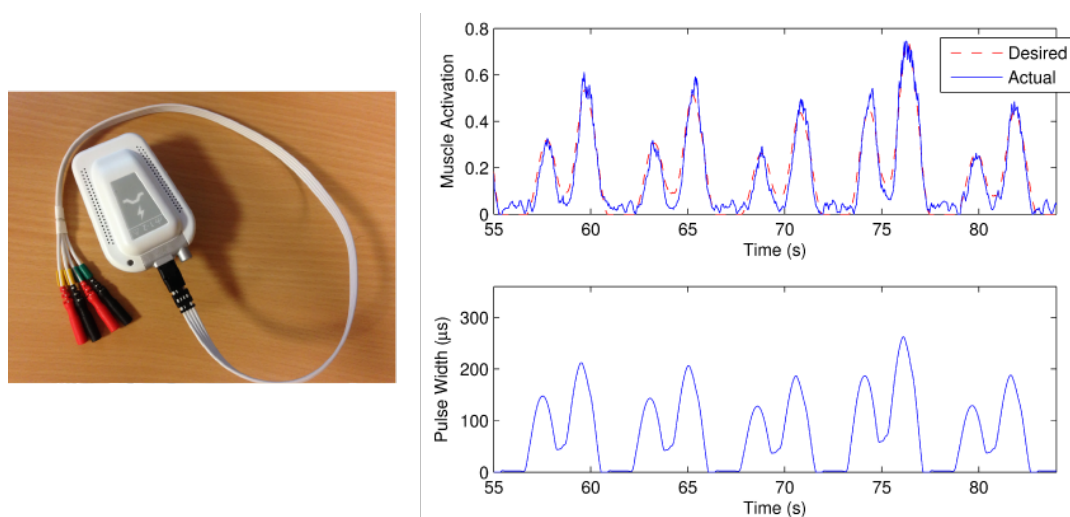


Figure 7. Left: Vivaltis portable stimulator; Right: Real-time control performance of muscle activation with desired dual sinusoidal shaped muscle activation pattern (red dash line is desired muscle activation trajectory and blue solid line is the measured muscle activation pattern under the muscle activation control by FES). The lower plot is the corresponding computed stimulation pulse width.

5.2.3. Direct spinal stimulation for rehabilitation of bladder, bowel and sexual functions in spinal cord injury

Participants: Christine Azevedo Coste, Luc Bauchet [CHU Montpellier], Claire Delleci [CHU Bordeaux], Charles Fattal [CRF DIVIO, Dijon, France], Thomas Guiho, David Guiraud, Jean-Rodolphe Vignes [CHU Bordeaux].

Complete spinal cord injury results in loss of movement and sensory sensations but also in function of organs. For example, nearly all spinal cord injured subjects lose their bladder control and are prone to kidney failure if they do not apply intermittent (self-) catheterization. Electrical stimulation of the sacral spinal roots with an implantable neuroprosthesis is one option besides self-catheterization to become continent and control micturition. However, many persons do not ask for this neuroprosthesis since deafferentation and loss of sensory functions and reflexes are serious side effects. Spinal cord stimulation (SCS) is a general term which includes both epidural and intradural stimulation. Originally associated with the treatment of chronic

neurological pain (in the 1970ies), SCS led also to immediate and profound improvements of sensory and motor functions in recent studies both on SCI patients (only on very few case studies) and rodents. Despite these promising results some limitations have still to be overcome. Among them, the use of small animal models, the empirical aspect of the stimulation procedure and the impact of these protocols on intestinal and urinary functions are critical. To counteract these limits, we want to explore intradural and epidural stimulations in an intermediate model- the house pig- and assess their impact on bladder, guts and genitals. In order to evaluate our approach, we will record EMG signals of lower limbs and sphincters (both urethral and anal), and simultaneously, we will monitor bladder and rectal pressure.

Already preliminary experimental explorations were performed with direct spinal cord stimulation in June (on 2 animals). Experiments were conducted under neurosurgeons involved in the project and urodynamics was recorded together with rectum pressure and sphincters EMG during each stimulation session.

5.3. Neuroprostheses and technology

5.3.1. Selectivity of nerve stimulation using a 12 pole multipolar cuff

Participants: Wafa Tigra, Olivier Rossel, Thomas Guiho, David Guiraud, Christine Azevedo Coste, Hubert Taillades [UM].

Experimentations were performed on 5 rabbits (New Zealand white). A multipolar cuff electrode (12 poles, diameter 3 mm, length 20 mm, 12 oblong contacts of 5mm length) was placed around the sciatic nerve of the rabbit 3 cm above the tibiofibular bifurcation. The nerve was stimulated with increasing intensity. The protocol consisted of the activation of one or more channels of the electrode, the input is a biphasic asymmetric stimulation and the pulse width is modulated in intensity (up to 2.4 mA) and fixed in length (250 μ sec, 100 μ sec interstim). A stimulus (4 Hz) is used for 2 seconds. 48 configurations of stimulation were tested. Needle electrodes were inserted on the lateral and medial gastrocnemius, soleus, tibialis and extensor digitorum muscles to record EMG signals and were used to evaluate the selectivity capacities of given cuff electrode configuration. The rabbit foot was also attached to a force platform. Inter-fascicular selectivity was observed for the 5 animals. Intra fascicular selectivity was also observed in 3 animals. Placed at a single location, our cuff electrode is capable to activate, selectively, some muscles. Experiments were performed under ethical committee agreement at the "Plateau Technique Chirurgie Experimentale" (Montpellier).

5.3.2. A novel EMG interface for individuals with quadriplegia to pilot robot hand grasping

Participants: Wafa Tigra, Benjamin Navarro [LIRMM], Andrea Cherubini [LIRMM], Xavier Gorrion [???], Anthony Gelis [PROPARGA], Charles Fattal [CRF DIVIO, Dijon, France], David Guiraud, Christine Azevedo Coste.

We have developed and validated a new human-machine interface dedicated to individuals with quadriplegia. We investigated the feasibility of online processing sus-lesional muscle responses, to pilot an assistive device. The ability to voluntarily contract a set of selected muscles was assessed in five spinal cord injured subjects through electromyography analysis. Two subjects have also been asked to use the EMG interface to control palmar and lateral grasping of a robot hand (fig.8). These preliminary results sound very promising and open the way to new interface solutions for high level spinal cord injured patients(fig.8).

5.3.3. Wearable 56-pole stimulator

Participants: Arthur Hiarrassary, David Andreu, David Guiraud, Olivier Rossel, Thomas Guiho.

In the context of the EPIONE European project, we have designed and developed, with Axonic, a wearable multichannel stimulator (fig.9), to face phantom limb pain (PLP). This 56-pole neural stimulator is based on four 16-pole stimulation units (each one being connected to a 16-pole TLIFE intra-fascicular electrode) connected to a real-time controller by means of an embedded deterministic network. This controller, in charge of executing FES functions (threshold determination, sensation characterization, etc.), pilots the stimulation units and allows for real-time modulation of the multisite stimulation. The controller can be remotely configured and exploited by the practitioner, by means of a dedicated software (Synergy Neuromodulation Software). But it has been also connected to the controller of the EPFL's artificial hand, in order to link hand touch sensors with neural stimulation to induce natural, meaningful sensations to the amputee.

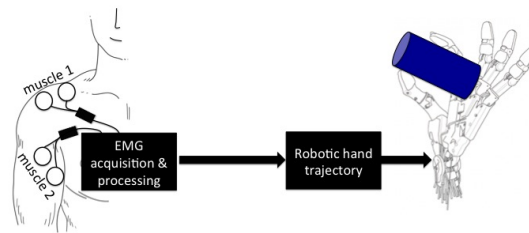


Figure 8. Principle of robot hand control through EMG signals; Setup description and upper arm positioning during EMG recordings.

This stimulator has been deeply validated through animal experiments (rats and pigs, respectively with UAB Barcelone and SMI Alboorg) and is currently used on human at UCBM Rome (<http://project-epione.eu/>).



Figure 9. 56-pole neural stimulator

5.3.4. CORAIL: Neural Stimulation Integrated Circuit

Participants: Jérémie Salles, Guy Cathébras, Milan Demarcq, David Guiraud, Guillaume Souquet, David Andreu.

DEMAR is currently finishing the development of CORAIL (Current Output Reconfigurable Asic Interface Low power), a new ASIC dedicated to electric neural stimulation.

Its main analog characteristics are:

- 12 independent current output channels;
- a full-scale current of 5 mA with a quantum of 1.3 μ A;
- a symmetrical power supply (\pm VHT with ground midpoint).

This front-end integrated circuit is designed to perform multipolar electrical stimulation of the nerve with highly configurable waveforms in order to achieve selective activation of organs or muscles. In comparison with previously developed current output ASICs, the CORAIL IC embeds new features such as the storage of multiple electrode configurations or the possibility to internally combine poles. These specific aspects of CORAIL and the fact that its elaboration benefited from clinical experience in the team will allow enhanced integration within the whole electrical stimulation environment.

The resulting stimulation ASIC aims to be part of an implanted distributed stimulation system, composed of multiple stimulation units spread across the body, in which CORAIL will be the front-end entity in charge of delivering the current to the electrode. Thus, special care has been paid to its integration in such a network with an emphasis on low power consumption for which different mechanisms have been implemented.

The ASIC is currently undergoing the last phases of its development and a first version is due for fabrication in February 2016.

5.3.5. Tele-Rehabilitation Platform for Gait Training

Participants: Mitsuhiro Hayashibe, Antonio P.I. Bo [Universidade de Brasilia, Brasil], Leslie Casas [Pontificia Universidad Catolica del Peru, Peru], Gonzalo Cucho [Pontificia Universidad Catolica del Peru, Peru], Dante Elias [Pontificia Universidad Catolica del Peru, Peru].

Throughout the world there is an increasing need for better technologies for rehabilitation and assistance. These new solutions must present improved performance in terms of therapy effectiveness, while at the same time minimizing the corresponding costs. In this scenario, computer-aided methods represent a promising alternative for the challenges currently faced by the rehabilitation domain. A tele-rehabilitation platform for gait training in intercontinental circumstances is developed under STIC-AmSud program. This project was joint program 2012-2013 among Inria France, UnB (University of Brasilia) and PCUP (Pontifical Catholic University of Peru) for tele-rehabilitation framework. This system has two mode: Self-modulation control in which the subject can control the speed of the motion therapy with his comfortable training speed and Guidance control mode in which the motion transfer is performed from one therapist to one patient. Guidance control can be performed both with local data transmission and intercontinental data transmission. Fig. 10 shows the case where the motion transfer regarding foot placement was performed with local data transmission. The test with intercontinental data transmission was also realized between France and Peru.



Figure 10. Tele-rehabilitation platform for gait training: Guidance control mode.

5.3.6. Control and scheduling co-design for stimulation systems

Participants: Daniel Simon, David Andreu.

Functional Electrical Stimulation (FES) is used in therapy for rehabilitation or substitution for disabled people. They are control systems using electrodes to interface a digital control system with livings. Hence the whole system gathers continuous-time (muscles and nerves) and discrete-time (controllers and communication links) components. During the design process, realistic simulation remains a precious tool ahead of real experiments to check without danger that the implementation matches the functional and safety requirements. To this aim we are developing a real-time open software simulation system, dedicated to the analysis of FES systems deployed over distributed execution resources and wireless links. The simulation tool is especially devoted to the joint design and analysis of control loops and real-time features.

Realistic simulations are effective tools to design and tune complex systems whose analysis cannot be provided only by theory. Several simulation steps can be explored, from simple functional analysis to HIL, to design, test, tune and validate both the single components of the system and their interactions in a distributed architecture. Simulations are precious, as they allow for non-destructive trials, which must be considered in any domain but this is of particular interest for bio-engineering [42].

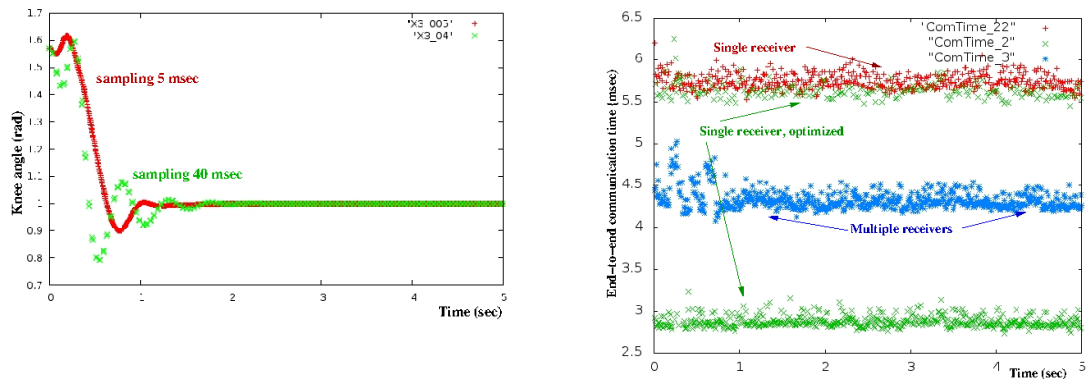


Figure 11. Simulation traces : knee position w.r.t. sampling and communication delays

It is expected that this particular simulator may provide inputs in two main directions. Firstly it allows for preliminary testing and tuning new FES protocols without needing for real experiments with patients, and may help for writing the ethical protocols needed for any experiments involving livings. Secondly it can be used to preliminary evaluation of new technologies or implementations, without costly reworking of existing electronic chips or certified components.

The simulation software is open, so that enhancements w.r.t. to the original release can be added upon request of various designers and to fulfill various objectives.

5.3.7. Control loops design principles for autonomic computing

Participants: Daniel Simon, Eric Rutten [Inria Grenoble Rhône-Alpes], Nicolas Marchand [GIPSA-lab].

Computing systems are becoming more and more dynamically reconfigurable or adaptive, to be flexible w.r.t. their environment and to automate their administration. Autonomic computing proposes a general structure of feedback loop to take this into account. We are particularly interested in approaches where this feedback loop is considered as a case of control loop where techniques stemming from Control Theory can be used to design efficient safe, and predictable controllers. This approach is emerging, with separate and dispersed effort, in different areas of the field of reconfigurable or adaptive computing, at software or architecture level.

We aim at conveying to Computer Scientists the interest and advantages of adopting a Control Theory perspective for the efficient and predictable design of autonomic systems. Compared with open-loop, closed-loop control provides adaptability and robustness, allowing for the design of fault-tolerant systems against varying and uncertain operating conditions. However, there still is a deep need for research in the problems of mapping from high-level objectives in terms of Quality of Service (QoS) or Service Level Objectives (SLO) and abstract models towards lower-levels effective actions on the managed systems [46].

5.4. Others

5.4.1. Do doors opening affect the air contamination in clean surgery? A Prospective, Cross-sectional Study (the ARIBO Project)

Participants: Gabriel Birgand [APHP], Christine Azevedo Coste, Stephane Rukly [INSERM], Roger Pissard-Gibollet [Inria Grenoble Rhône-Alpes], Jean-Christophe Lucet [APHP].

Inappropriate staff behaviours can lead to environmental contamination in the operating room (OR) and subsequent surgical site infection (SSI). This study focused on the continued assessment of OR staff behaviours using doors sensors and their impact on the SSI risk during surgical procedures. This multicentre observational study included 13 ORs in 10 hospitals, 5 University hospitals and 5 private hospitals. Two specialties of clean surgery with cutaneous approach were included: cardiac surgery with procedures requiring a full median sternotomy (CABG or valve replacement surgery); and planned orthopaedic surgery for total hip (THR) or knee replacement (TKR). For each surgical specialty involved, the observed ORs were randomly selected. Doors opening were observed by means of wireless inertial sensors fixed on the doors. For each surgical procedure, 3 microbiological air counts, continuous particles counts of 0.3, 0.5 and $5\mu\text{m}$ particles, and one bacteriological sample of the wound before skin closure were performed. We collected informations on the OR staff, surgical procedures and surgical environment characteristics. Statistics were performed using univariate and multivariate analysis to adjust on aerolic and architectural characteristics of the OR. We included 34 orthopaedic and 26 cardiac procedures. The mean duration of intervention, from patient entry to exit in the OR, was 5.3 (SD 1.1) h. in cardiac and 2.6 (0.7) h. in orthopaedic surgery. The median number of doors opening was 146 (IQR: 121-183; Min-Max: 86-319) per intervention and 29 (IQR: 23-36; Min-Max: 17, 54) per h. in cardiac surgery and 71.5 (IQR: 58-92; Min-Max: 54-136) per intervention and 29 (IQR: 25-34; Min-Max: 16-65) per h. in orthopaedic procedures. Doors stayed open in average 43 minutes (Min-Max: 19-115) in cardiac and 36 (8-199) in orthopaedic, representing respectively 13.5 percents and 23 percents of the duration of intervention. The highest frequency of doors opening was observed between wound closure and patient exit, median 20.1 openings/h (12.5-32.3) and from patient entry to the incision 13.2 openings/h (8-19). The number and duration of doors opening was significantly different between centres (higher in university hospital, $p < 0.01$). High frequency of openings was observed for doors that should stay closed during procedures (materials store, decontamination room). The number of doors opening from skin incision to wound closure affected significantly the 0.5 and $5\mu\text{m}$ particles count ($p < 0.01$ and 0.02 respectively). This study based on automatic observation suggests a large heterogeneity of doors openings between types of interventions, ORs and hospitals. Data give a standard of doors opening for CABG, THR and TKR. Door openings affected air contamination, potentially jeopardizing operating room sterility. The causes and influences of behaviours in the OR must be evaluated to identify ways to reduce the associated risks.

DRACULA Project-Team

6. New Results

6.1. Implication of the autologous immune system in BCR-ABL transcript variations in chronic myelogenous leukemia patients treated with Imatinib

Imatinib (IM) and other tyrosine kinase inhibitors (TKI) have improved treatment of chronic myelogenous leukemia (CML); however, most patients are not cured. Deeper mechanistic understanding may improve TKI combination therapies to better control the residual leukemic cell population. In analyzing our patients' data, we found that many patients who otherwise responded well to IM therapy still showed variations in their BCR-ABL transcripts. To investigate this phenomenon, we applied a mathematical model (see [14]) that integrates CML and an autologous immune response to the patients' data. We define an immune window, or a range of leukemic loads for which the autologous immune system induces an improved response. Our modeling results in [14], suggest that, at diagnosis, a patient's leukemic load is able to partially or fully suppress the autologous immune response developed in a majority of patients, towards the CML clone(s). IM therapy drives the leukemic population into the "immune window", allowing the patient's autologous immune cells to expand and eventually mount an efficient recognition of the residual leukemic burden. This response drives the leukemic load below this immune window, allowing the leukemic population to partially recover until another weaker immune response is initiated. Thus, the autologous immune response may explain the oscillations in BCR-ABL transcripts regularly observed in patients on IM.

6.2. Predicting pathogen-specific CD8 T cell immune responses from a modeling approach

The primary CD8 T cell immune response constitutes a major mechanism to fight an infection by intra-cellular pathogens. We aim at assessing whether pathogen-specific dynamical parameters of the CD8 T cell response can be identified, based on measurements of CD8 T cell counts, using a modeling approach. We generated experimental data consisting in CD8 T cell counts kinetics during the response to three different live intra-cellular pathogens: two viruses (influenza, vaccinia) injected intranasally, and one bacteria (*Listeria monocytogenes*) injected intravenously. All pathogens harbor the same antigen (NP68), but differ in their interaction with the host. In parallel, we developed in [16] a mathematical model describing the evolution of CD8 T cell counts and pathogen amount during an immune response. This model is characterized by 9 parameters and includes relevant feedback controls. The model outputs were compared with the three data series and an exhaustive estimation of the parameter values was performed. By focusing on the ability of the model to fit experimental data and to produce a CD8 T cell population mainly composed of memory cells at the end of the response, critical parameters were identified. We show that a small number of parameters (2 – 4) define the main features of the CD8 T cell immune response and are characteristic of a given pathogen. Among these parameters, two are related to the effector CD8 T cell mediated control of cell and pathogen death. The parameter associated with memory cell death is shown to play no relevant role during the main phases of the CD8 T cell response, yet it becomes essential when looking at the predictions of the model several months after the infection.

6.3. Dynamics of cell generation and turnover in the human heart

The contribution of cell generation to physiological heart growth and maintenance in humans has been difficult to establish and has remained controversial. We report in [8] that the full complement of cardiomyocytes is established perinatally and remains stable over the human lifespan, whereas the numbers of both endothelial and mesenchymal cells increase substantially from birth to early adulthood. Analysis of the integration of

nuclear bomb test-derived ^{14}C revealed a high turnover rate of endothelial cells throughout life ($> 15\%$ per year) and more limited renewal of mesenchymal cells ($< 4\%$ per year in adulthood). Cardiomyocyte exchange is highest in early childhood and decreases gradually throughout life to $< 1\%$ per year in adulthood, with similar turnover rates in the major subdivisions of the myocardium. We provide an integrated model of cell generation and turnover in the human heart.

6.4. Travelling waves of cell differentiation

The paper [7] is devoted to modelling of cell differentiation in an initially homogeneous cell population. The mechanism which provides coexistence of two cell lineages in the initially homogeneous cell population is suggested. If cell differentiation is initiated locally in space in the population of undifferentiated cells, it can propagate as a travelling wave converting undifferentiated cells into differentiated ones. We suggest a model of this process which takes into account intracellular regulation, extracellular regulation and different cell types. They include undifferentiated cells and two types of differentiated cells. When a cell differentiates, its choice between two types of differentiated cells is determined by the concentrations of intracellular proteins. Differentiated cells can either stimulate differentiation into their own cell lineage or into another cell lineage. In the case of the positive feedback, only one lineage of differentiated cells will finally appear. In the case of negative feedback, both of them can coexist. In this case a periodic spatial pattern emerges behind the wave.

6.5. Pattern regeneration based on cell memory

In [24], we present a new model of the cellular dynamics that enable regeneration of complex biological morphologies. Biological cell structures are considered as an ensemble of mathematical points on the plane. Each cell produces a signal which propagates in space and is received by other cells. The total signal received by each cell forms a signal distribution defined on the cell structure. This distribution characterizes the geometry of the cell structure. If a part of this structure is removed, the remaining cells have two signals. They keep the value of the signal which they had before the amputation (memory), and they receive a new signal produced after the amputation. Regeneration of the cell structure is stimulated by the difference between the old and the new signals. It is stopped when the two signals coincide. The algorithm of regeneration contains certain rules which are essential for its functioning, being the first quantitative model of cellular memory that implements regeneration of complex patterns to a specific target morphology. Correct regeneration depends on the form and the size of the cell structure, as well as on some parameters of regeneration.

6.6. Target morphology and cell memory

Despite the growing body of work on molecular components required for regenerative repair, we still lack a deep understanding of the ability of some animal species to regenerate their appropriate complex anatomical structure following damage. A key question is how regenerating systems know when to stop growth and remodeling – what mechanisms implement recognition of correct morphology that signals a stop condition? In [11], we review two conceptual models of pattern regeneration that implement a kind of pattern memory. In the first one, all cells communicate with each other and keep the value of the total signal received from the other cells. If a part of the pattern is amputated, the signal distribution changes. The difference from the original signal distribution stimulates cell proliferation and leads to pattern regeneration, in effect implementing an error minimization process that uses signaling memory to achieve pattern correction. In the second model, we consider a more complex pattern organization with different cell types. Each tissue contains a central (coordinator) cell that controls the tissue and communicates with the other central cells. Each of them keeps memory about the signals received from other central cells. The values of these signals depend on the mutual cell location, and the memory allows regeneration of the structure when it is modified. The purpose of these models is to suggest possible mechanisms of pattern regeneration operating on the basis of cell memory which are compatible with diverse molecular implementation mechanisms within specific organisms.

6.7. Transplanted bone marrow-derived cells contribute to human adipogenesis

Because human white adipocytes display a high turnover throughout adulthood, a continuous supply of precursor cells is required to maintain adipogenesis. Bone marrow (BM)-derived progenitor cells may contribute to mammalian adipogenesis; however, results in animal models are conflicting. In [22], we demonstrate in 65 subjects who underwent allogeneic BM or peripheral blood stem cell (PBSC) transplantation that, over the entire lifespan, BM/PBSC-derived progenitor cells contribute 10% to the subcutaneous adipocyte population. While this is independent of gender, age, and different transplantation-related parameters, body fat mass exerts a strong influence, with up to 2.5-fold increased donor cell contribution in obese individuals. Exome and whole-genome sequencing of single adipocytes suggests that BM/PBSC-derived progenitors contribute to adipose tissue via both differentiation and cell fusion. Thus, at least in the setting of transplantation, BM serves as a reservoir for adipocyte progenitors, particularly in obese subjects.

6.8. Modelling of platelet–fibrin clot formation in flow

The paper [23] is devoted to mathematical modelling of clot growth in blood flow. Great complexity of the hemostatic system dictates the need of usage of the mathematical models to understand its functioning in the normal and especially in pathological situations. In this work we investigate the interaction of blood flow, platelet aggregation and plasma coagulation. We develop a hybrid DPD–PDE model where dissipative particle dynamics (DPD) is used to model plasma flow and platelets, while the regulatory network of plasma coagulation is described by a system of partial differential equations. Modelling results confirm the potency of the scenario of clot growth where at the first stage of clot formation platelets form an aggregate due to weak inter-platelet connections and then due to their activation. This enables the formation of the fibrin net in the centre of the platelet aggregate where the flow velocity is significantly reduced. The fibrin net reinforces the clot and allows its further growth. When the clot becomes sufficiently large, it stops growing due to the narrowed vessel and the increase of flow shear rate at the surface of the clot. Its outer part is detached by the flow revealing the inner part covered by fibrin. This fibrin cap does not allow new platelets to attach at the high shear rate, and the clot stops growing. Dependence of the final clot size on wall shear rate and on other parameters is studied.

6.9. Conceptual model of morphogenesis and regeneration

The paper [24] is devoted to computer modelling of the development and regeneration of multicellular biological structures. Some species (e.g. planaria and salamanders) are able to regenerate parts of their body after amputation damage, but the global rules governing cooperative cell behaviour during morphogenesis are not known. Here, we consider a simplified model organism, which consists of tissues formed around special cells that can be interpreted as stem cells. We assume that stem cells communicate with each other by a set of signals, and that the values of these signals depend on the distance between cells. Thus the signal distribution characterizes location of stem cells. If the signal distribution is changed, then the difference between the initial and the current signal distribution affects the behaviour of stem cells—e.g. as a result of an amputation of a part of tissue the signal distribution changes which stimulates stem cells to migrate to new locations, appropriate for regeneration of the proper pattern. Moreover, as stem cells divide and form tissues around them, they control the form and the size of regenerating tissues. This two-level organization of the model organism, with global regulation of stem cells and local regulation of tissues, allows its reproducible development and regeneration.

6.10. Delay differential-difference system for hematopoietic stem cell dynamics

We investigate in [2] and [3] a mathematical model of hematopoietic stem cell dynamics. We take two cell populations into account, quiescent and proliferating one, and we note the difference between dividing cells that enter directly to the quiescent phase and dividing cells that return to the proliferating phase to divide again. The resulting mathematical model is a system of two age-structured partial differential equations. By

integrating this system over age and using the characteristics method, we reduce it to a delay differential-difference system, and we investigate the existence and stability of the steady states. We give sufficient conditions for boundedness and unboundedness properties for the solutions of this system. By constructing a Lyapunov function, the trivial steady state, describing cell's dying out, is proven to be globally asymptotically stable when it is the only equilibrium. The stability analysis of the unique positive steady state, the most biologically meaningful one, and the existence of a Hopf bifurcation allow the determination of a stability area, which is related to a delay-dependent characteristic equation. Numerical simulations illustrate our results on the asymptotic behavior of the steady states and show very rich dynamics of this model. This study may be helpful in understanding the uncontrolled proliferation of blood cells in some hematological disorders.

6.11. Discrete limit and monotonicity properties of the Floquet eigenvalue in an age structured cell division cycle model

We consider in [19] a cell population described by an age-structured partial differential equation with time periodic coefficients. We assume that division only occurs after a minimal age (majority) and within certain time intervals. We study the asymptotic behavior of the dominant Floquet eigenvalue, or Perron-Frobenius eigenvalue, representing the growth rate, as a function of the majority age, when the division rate tends to infinity (divisions become instantaneous). We show that the dominant Floquet eigenvalue converges to a staircase function with an infinite number of steps, determined by a discrete dynamical system. As an intermediate result, we give a structural condition which guarantees that the dominant Floquet eigenvalue is a nondecreasing function of the division rate. We also give a counter example showing that the latter monotonicity property does not hold in general.

6.12. Optimal linear stability condition for scalar differential equations with distributed delay

Linear scalar differential equations with distributed delays appear in the study of the local stability of nonlinear differential equations with feedback, which are common in biology and physics. Negative feedback loops tend to promote oscillations around steady states, and their stability depends on the particular shape of the delay distribution. Since in applications the mean delay is often the only reliable information available about the distribution, it is desirable to find conditions for stability that are independent from the shape of the distribution. We show in [9] that for a given mean delay, the linear equation with distributed delay is asymptotically stable if the associated differential equation with a discrete delay is asymptotically stable. We illustrate this criterion on a compartment model of hematopoietic cell dynamics to obtain sufficient conditions for stability.

6.13. A mathematical model of leptin resistance

Obesity is often associated with leptin resistance, which leads to a physiological system with high leptin concentration but unable to respond to leptin signals and to regulate food intake. We propose in [20] a mathematical model of the leptin-leptin receptors system, based on the assumption that leptin is a regulator of its own receptor activity, and investigate its qualitative behavior. Based on current knowledge and previous models developed for body weight dynamics in rodents, the model includes the dynamics of leptin, leptin receptors and the regulation of food intake and body weight. It displays two stable equilibria, one representing a healthy state and the other one an obese and leptin resistant state. We show that a constant leptin injection can lead to leptin resistance and that a temporal variation in some parameter values influencing food intake can induce a change of equilibrium and a pathway to leptin resistance and obesity.

DYLISS Project-Team

7. New Results

7.1. Data integration

Participants: Jacques Nicolas, Charles Bettembourg, Jérémie Bourdon, Jeanne Got, Marie Chevallier, Guillaume Collet, Olivier Dameron, Damien Eveillard, Julie Laniau, Anne Siegel.

Extended notions of sign consistency to relate experimental data to signaling and regulatory network topologies. Interaction graphs provide a suitable representation of cellular networks with information flows. Methods based on sign consistency have been shown to be valuable tools to (i) predict qualitative responses, (ii) test the consistency of network topologies and experimental data, and (iii) apply repair operations to the network model suggesting missing or wrong interactions. We present a framework to unify different notions of sign consistency and propose a refined method for data discretization that considers uncertainties in experimental profiles. We furthermore introduce a new constraint to filter undesired model behaviors induced by positive feedback loops. Finally, we generalize the way predictions can be made by the sign consistency approach. This corresponds to an extension of our *Bioquali* software. [Anne Siegel] [21]

Putative bacterial interactions from metagenomic knowledge with an integrative systems ecology approach. Our software tool *shogen* was used to decipher functional roles within a consortium of five mining bacteria through the integration of genomic and metabolic knowledge at genome scale. We first reconstructed a global metabolic network. Next, using a parsimony assumption, we deciphered sets of genes, called Sets from Genome Segments (SGS), that (i) are close on their respective genomes, (ii) take an active part in metabolic pathways and (iii) whose associated metabolic reactions are also closely connected within metabolic networks. The use of SGS (*shogen*) pinpoints a functional compartmentalization among the investigated species and exhibits putative bacterial interactions necessary for promoting these pathways. [Damien Eveillard, Anne Siegel] [17]

Optimal Threshold Determination for Interpreting Semantic Similarity and Particularity: Application to the Comparison of Gene Sets and Metabolic Pathways Using GO and ChEBI. We developed a method for determining optimal semantic similarity and particularity thresholds in order to interpret the results of the comparison of ontology terms sets. We applied this method on the GO and ChEBI ontologies. Qualitative analysis using the thresholds on the PPAR multigene family yielded biologically-relevant patterns. [Charles Bettembourg, Olivier Dameron] [16]

AskOmics : Integration et interrogation de réseaux de régulation génomique et post-génomique. We present AskOmics, an integration and interrogation software using a RDF model and the SPARQL query language. The purpose of this work is to obtain quick answers to biological questions demanding currently hours of manual search in several spreadsheet results files. AskOmics allows biologists to integrate and interrogate their data by themselves without any knowledge about RDF and SPARQL required. [Charles Bettembourg, Olivier Dameron] [30]

7.2. Time-series and asymptotic dynamics

Participants: Anne Siegel, Jacques Nicolas, Jérémie Bourdon, Jean Coquet, Victorien Delannée, Vincent Picard, Nathalie Théret.

Identification of logical models for signaling pathways: towards a systems biology loop. Logical models of signaling pathways are a promising way of building effective in silico functional models of a cell. The automated learning of Boolean logic models describing signaling pathways can be achieved by training to phosphoproteomics data. This data is unavoidably subject to noise. As a result, the learning process leads to a family of feasible logical networks rather than a single model. This family is composed of logic models proposing different internal wirings for the system, implying that the logical predictions from this family may suffer a significant level of variability leading to uncertainty. In our work, combinatorial optimization methods based on recent logic programming paradigm allow to enumerate, and discriminate the family of logical models explaining data. Together, these approaches enable a robust understanding of the system response. The results are implemented in the *caspo* software [Jacques Nicolas, Anne Siegel] [22], [23]

Boolean Network Identification from Multiplex Time Series Data. The ASP-based learning algorithm developed in the team to train logical models of signaling networks focuses on the comparison of two time-points and assumes that the system has reached an early steady state. We have generalized such a learning procedure in order to discriminate Boolean networks according to their transient dynamics. To that goal, we exhibit a necessary condition that must be satisfied by a Boolean network dynamics to be consistent with a discretized time series trace. This approach was included in the ASP-based framework designed for the *caspo* software. We ended up with a global learning algorithm and compared it to learning approaches based on static data. [Anne Siegel] [31]

Representation of symbolic dynamical systems generated by a substitution. Iterated morphisms are combinatorial processes which are related to several classes of dynamical systems appearing in several fields of computer sciences and mathematics: numeration, ergodic theory, discrete geometry. They may be associated to fractal sets called "Rauzy fractals" whose topological properties are linked to the properties of the underlying dynamical system. We have introduced a generic algorithm framework to check such topological properties within a complete family of iterated morphism. This makes efficient the verification of conjectures on several families of substitutions related to multi-dimensional continued fraction algorithms. [Anne Siegel] [32], [25], [14]

Multivariate Normal Approximation for the Stochastic Simulation Algorithm: Limit Theorem and Applications. We present a central limit theorem for the Gillespie stochastic trajectories when the living system has reached a steady-state, that is when the internal bio-molecules concentrations are assumed to be at equilibrium. It appears that the stochastic behavior in steady-state is entirely characterized by the stoichiometry matrix of the system and a single vector of reaction probabilities. We propose several applications of this result such as deriving multivariate confidence regions for the time course of the system and a constraints-based approach which extends the flux balance analysis framework to the stochastic case. [J er mie Bourdon, Vincent Picard, Anne Siegel] [20], [12]

A Logic for Checking the Probabilistic Steady-State Properties of Reaction Networks. Designing probabilistic reaction models and determining their stochastic kinetic parameters are major issues in systems biology. In order to assist in the construction of reaction network models, we introduce a logic that allows one to express asymptotic properties about the steady-state stochastic dynamics of a reaction network. Basically, the formulas can express properties on expectancies, variances and co-variances. We demonstrate that deciding the satisfiability of a formula is NP-hard. [J er mie Bourdon, Vincent Picard, Anne Siegel] [28], [12]

7.3. Sequence and structure annotation

Participants: Fran ois Coste, Aymeric Antoine-Lorquin, Catherine Belleann e, Guillaume Collet, Clovis Galiez, Laurent Miclet, Jacques Nicolas.

Amplitude Spectrum Distance: measuring the global shape divergence of protein fragments. We introduce here the Amplitude Spectrum Distance (ASD), a novel way of comparing protein fragments based on the discrete Fourier transform of their C_α distance matrix. Defined as the distance between their amplitude spectra, ASD can be computed efficiently and provides a parameter-free measure of the global shape dissimilarity of two fragments. ASD inherits from nice theoretical properties, making it tolerant to shifts, insertions,

deletions, circular permutations or sequence reversals while satisfying the triangle inequality. The practical interest of ASD with respect to RMSD, RMSD_d, BC and TM scores is illustrated through zinc finger retrieval experiments and concrete structure examples. The benefits of ASD are also illustrated by two additional clustering experiments: domain linkers fragments and complementarity-determining regions of antibodies. [Clovis Galiez, François Coste] [19]

Structural conservation of remote homologues: better and further in contact fragments. We address a basic question on sequence-structure relationships in proteins: does a protein sequence depict a structure with a uniform faithfulness all along the sequence? We investigate this question by defining contact fragments. This study suggests that sequence homologs of CF are significantly more faithful to structure than randomly chosen fragments, so that CF carry a strong sequence-structure relationship, allowing them to be used as accurate building blocks for structure prediction. [Clovis Galiez, François Coste] [26]

VIRALpro: a tool to identify viral capsid and tail sequences. Not only sequence data continues to outpace annotation information, but the problem is further exacerbated when organisms are underrepresented in the annotation databases. This is the case with non human-pathogenic viruses which occur frequently in metagenomic projects. Thus there is a need for tools capable of detecting and classifying viral sequences. We describe VIRALpro a new effective tool for identifying capsid and tail protein sequences, which are the cornerstones toward viral sequence annotation and viral genome classification. [Clovis Galiez, François Coste] [18]

Finding Optimal Discretization Orders for Molecular Distance Geometry. The Molecular Distance Geometry Problem (MDGP) is the problem of finding the possible conformations of a molecule by exploiting available information about distances between some atom pairs. Under minimal assumptions the MDGP can be discretized so that the search domain of the problem becomes a tree that can be explored by using an interval Branch & Prune (iBP) algorithm. In this context, the discretization assumptions are strongly dependent on the atomic ordering, which can also impact the computational cost of the iBP algorithm. In this work, we propose a new partial discretization order for protein backbones. This new atomic order optimizes a set of objectives that aim at improving the iBP performances. The optimization of the objectives is performed by Answer Set Programming (ASP), which allows to express the problem by a set of logical constraints. The comparison with previously proposed orders for protein backbones shows that this new discretization order makes iBP perform more efficiently. [Jacques Nicolas] [34]

From formal concepts to analogical complexes. Reasoning by analogy is an important component of common sense reasoning whose formalization has undergone recent improvements with the logical and algebraic study of the analogical proportion. The starting point of this study considers analogical proportions on a formal context. We introduce analogical complexes, a companion of formal concepts formed by using analogy between four subsets of objects in place of the initial binary relation. They represent subsets of objects and attributes that share a maximal analogical relation. We show that the set of all complexes can be structured in an analogical complex lattice and give explicit formulae for the computation of their infimum and supremum. [Laurent Miclet, Jacques Nicolas] [27]

Comparison of the targets obtained by a scoring matrix and by a regular expression. Application to the search for LXR binding sites. In bioinformatics, it is a common task to search for new instances of a pattern built from a set of reference sequences. For the simplest and most frequent cases, patterns are represented in two ways: regular expression or scoring matrix. Since both representations seem to be used indifferently in practice, one may wonder if they have any impact on the result. This study compares hits obtained with scoring matrices or by regular expressions allowing up to two substitutions. It shows that, in our LXR study, sequences found by a scoring matrix are closer to the targeted hits than sequences found by a regular expression. [Aymeric Antoine-Lorquin, Jacques Nicolas, Catherine Belleannée] [29]

Finding and Characterizing Repeats in Plant Genomes. Plant genomes contain a particularly high proportion of repeated structures of various types. This chapter proposes a guided tour of available softwares that can help biologists to look for these repeats and check some hypothetical models intended to characterize their structures. Since transposable elements are a major source of repeats in plants, we have provided a whole section on this topic as well as a selection of the main existing softwares. In order to better understand how

they work and how repeats may be efficiently found in genomes, the rest of the chapter is devoted to the foundations of the search for repeats and more complex patterns. We first introduce the key concepts that are useful for understanding the current state of the art in playing with words, applied to genomic sequences. In fact, biologists need to represent more complex entities where a repeat family is built on more abstract structures, including direct or inverted small repeats, motifs, composition constraints as well as ordering and distance constraints between these elementary blocks. The last section introduces concepts and practical tools that can be used to reach this syntactic level in biological sequence analysis. [*Jacques Nicolas*] [35]

ERABLE Project-Team

6. New Results

6.1. General comments

We present in this section the main results obtained in 2015. Some were already in preparation or submitted at the end of 2014. It will be indicated whenever this is the case.

We tried to organise the results following four of the five main axes of research of the team. Clearly, in some cases, a result obtained overlaps more than one axis. We chose the one that could be seen as the main concerned by such results. As concerns the Axis “Going towards control”, a work is in preparation that fits it. It will be presented in 2016.

We did not indicate here the results on more theoretical aspects of computer science if it did not seem for now that they could be relevant in contexts related to computational biology. Actually, we do believe those on scheduling (by, among others, A. Marchetti-Spaccamela and/or L. Stougie) [3], [38], [39], [10], [31], [23], [44], [43] or even one result related to context-free grammars (by, among others, P. Crescenzi) [11] could in the future become relevant for the life sciences (biology or ecology). However, we preferred for now to only indicate the theoretical results related to problems closely resembling questions that have already been addressed by us in computational biology.

Notice that such CS results concern not only cross-fertilising issues among different computational approaches, and we therefore extended the title of this axis for the purpose of presenting such results, for now purely theoretical.

A few other results are not mentioned either, not because the corresponding work is not important, but because it was likewise more specialised, or the work represented a survey.

6.2. Identifying the molecular elements

Genomic / NGS data management

Next-generation sequencing (NGS) technology has led the life sciences into the big data era. Today, sequencing genomes takes little time and cost, but yields terabytes of data to be stored and analysed. The biologists are often exposed to excessively time consuming and error-prone data management and analysis hurdles. We therefore proposed a database management system (DBMS) based approach to accelerate and substantially simplify genome sequence analysis [9]. To that aim, we extended MONETDB, an open-source column-based DBMS ([urlhttps://www.monetdb.org](https://www.monetdb.org)), with a BAM module, which enables easy, flexible, and rapid management and analysis of sequence alignment data stored as Sequence Alignment/Map (SAM/BAM) files. The main features of MONETDB/BAM were described using a case study on Ebola virus.

We also designed and realised a knowledge base for collecting, elaborating, and extracting analytical results of genomic, proteomic, biochemical, morphological investigations from animal models of cerebral stroke [45]. Data analysis techniques are tailored to make the data available for processing and correlation, in order to increase the predictive value of the preclinical data, to perform bio-simulation studies, and to support both academic and industrial research in the area of cerebral stroke therapy. The low reliability of animal models in replicating the human disease is one of the most serious problems in the field of medical and pharmaceutical research about stroke. The standard models for the study of ischaemic stroke are often poorly predictive as they simulate only partially the human disease. This work aims therefore at investigating animal models with diseases typically associated with the onset of stroke in human patients. A first statistical analysis of the retrieved information led to the validation of our animal models and suggested a predictive and translational value for parameters related to a specific model. In particular, concerning gene expression data, we applied a data analysis pipeline that initially takes into account an initial set of 64,000 genes and brought down the focus on a few tens of them.

NGS data analysis

The problem of enumerating bubbles with length constraints in directed graphs arises in transcriptomics where the question is to identify all alternative splicing events present in a sample of mRNAs sequenced by RNA-seq. We presented a new algorithm for enumerating bubbles with length constraints in weighted directed graphs [30]. This is the first polynomial delay algorithm for this problem and we showed that in practice, it is faster than previous approaches. This settled one of the main open questions from previous literature. Moreover, the new algorithm allows us to deal with larger instances and possibly detect longer alternative splicing events.

We also developed CIDANE, a novel framework for genome-based transcript reconstruction and quantification from RNA-seq reads [37]. CIDANE assembles transcripts with significantly higher sensitivity and precision than existing tools, while competing in speed with the fastest methods. In addition to reconstructing transcripts *ab initio*, the algorithm also allows to make use of the growing annotation of known splice sites, transcription start and end sites, or full-length transcripts, which are available for most model organisms. CIDANE supports the integrated analysis of RNA-seq and additional gene-boundary data and recovers splice junctions that are invisible to other methods.

SNPs (Single Nucleotide Polymorphisms) are genetic markers used in many areas of biology. Their precise identification is a prerequisite for association studies, which associate genotypes to phenotypes. Methods are currently developed for model species, but rely on the availability of a (good) reference genome, and cannot be applied to non-model species. They are also mostly tailored for whole genome (re-)sequencing experiments, whereas in many cases, transcriptome sequencing can be used as a cheaper alternative which already enables to identify SNPs located in transcribed regions. We proposed a method that identifies, quantifies and annotates SNPs without any reference genome, using RNA-seq data only. Individuals can be pooled prior to sequencing, if not enough material is available for sequencing from one individual. This pooling strategy still enables to allelotype loci and to associate them to phenotypes. Using human RNA-seq data, we first compared the performance of our algorithm, KISSPLICE, with GATK, a well established method that requires a reference genome. We showed that both methods perform similarly in terms of precision and recall. We then validated experimentally the predictions of our method using RNA-seq data from two non-model species. The method can be used for any species to annotate SNPs and to predict their impact on proteins. It can further be used to assess variants that are associated to a particular phenotype within a population, when replicates are provided for each biological condition. This work was submitted at the end of 2015.

Sequence alignment (full genomes or NGS data)

Sequence comparison is a fundamental step in many important tasks related to biology. Traditional algorithms for measuring approximation in sequence comparison are based on the notions of distance or similarity, and are generally computed through sequence alignment techniques. As circular genome structure is a common phenomenon in nature, a caveat of specialised alignment techniques for circular sequence comparison is that they are computationally expensive, requiring from super-quadratic to cubic time in the length of the sequences. We introduced a new distance measure based on q -grams, and showed how it can be computed efficiently for circular sequence comparison [41]. Experimental results, using real and synthetic data, demonstrated orders-of-magnitude superiority of our approach in terms of efficiency, while maintaining an accuracy very competitive to the state of the art.

Burrows-Wheeler Transform (BWT) has been successfully used to reduce the memory requirement for sequence alignment. We improved on previous results related to the problem of computing the Burrows-Wheeler Transform (BWT) using small additional space [12]. Our in-place algorithm does not need the explicit storage for the suffix sort array and the output array, as typically required in such previous work. It relies on the combinatorial properties of the BWT, and runs in $O(n^2)$ time in the comparison model using $O(1)$ extra memory cells, apart from the array of n cells storing the n characters of the input text. We then discussed the time-space trade-off when $O(k\sigma k)$ extra memory cells are allowed with σk distinct characters, providing an $O((n^2/k + n) \log^? k)$ -time algorithm to obtain (and invert) the BWT. In real systems where the alphabet size is a constant, for any arbitrarily small $\epsilon > 0$, the BWT of a text of n bytes can be computed in $O(n\sigma^{-1} \log n)$ time using just σn extra bytes.

Genome assembly problems

The human genome is diploid, which requires assigning heterozygous single nucleotide polymorphisms (SNPs) to the two copies of the genome. The resulting haplotypes, lists of SNPs belonging to each copy, are crucial for downstream analyses in population genetics. Currently, statistical approaches, which are oblivious to direct read information, constitute the state-of-the-art. Haplotype assembly, which addresses phasing directly from sequencing reads, suffers from the fact that sequencing reads of the current generation are too short to serve the purposes of genome-wide phasing. While future-technology sequencing reads will contain sufficient amounts of SNPs per read for phasing, they are also likely to suffer from higher sequencing error rates. Currently, no haplotype assembly approaches exist that allow for taking both increasing read length and sequencing error information into account. We developed WHATSHAP, the first approach that yields provably optimal solutions to the weighted minimum error correction problem in runtime linear in the number of SNPs [25]. WHATSHAP is a fixed parameter tractable (FPT) approach with coverage as the parameter. We demonstrated that WHATSHAP can handle datasets of coverage up to 20x, and that 15x are generally enough for reliably phasing long reads, even at significantly elevated sequencing error rates. We also find that the switch and flip error rates of the haplotypes we output are favourable when comparing them with state-of-the-art statistical phasers. By using novel combinatorial properties of Minimum Error Correction (MEC) instances, we were then able to provide new results on the fixed-parameter tractability and approximability of MEC [35]. In particular, we showed that MEC is in FPT when parameterised by the number of corrections, and, on “gapless” instances, it is in FPT also when parameterised by the length of the fragments, whereas the result known in the literature forces the reconstruction of complementary haplotypes. We then showed that MEC cannot be approximated within any constant factor while it is approximable within factor $O(\log nm)$ where nm is the size of the input. Finally, we provided a practical 2-approximation algorithm for the Binary MEC, a variant of MEC that has been applied in the framework of clustering binary data. Finally, by exploiting a feature of future-generation technologies – the uniform distribution of sequencing errors – we designed an exact algorithm, called HAPCOL, that is exponential in the maximum number of corrections for each SNP position and that minimises the overall error-correction score [26]. We performed an experimental analysis, comparing HAPCOL with the current state-of-the-art combinatorial methods both on real and simulated data. On a standard benchmark of real data, we showed that HAPCOL is competitive with state-of-the-art methods, improving the accuracy and the number of phased positions. Furthermore, experiments on realistically-simulated datasets revealed that HAPCOL requires significantly less computing resources, especially memory. Thanks to its computational efficiency, HAPCOL can overcome the limits of previous approaches, allowing to phase datasets with higher coverage and without the traditional all-heterozygous assumption.

Completing the genome sequence of an organism is an important task in comparative, functional and structural genomics. However, this remains a challenging issue from both a computational and an experimental viewpoint. Genome scaffolding (*i.e.* the process of ordering and orientating contigs) of *de novo* assemblies usually represents the first step in most genome finishing pipelines. We developed MEDUSA (Multi-Draft based Scaffold), an algorithm for genome scaffolding [6]. MEDUSA exploits information obtained from a set of (draft or closed) genomes from related organisms to determine the correct order and orientation of the contigs. MEDUSA formalises the scaffolding problem by means of a combinatorial optimisation formulation on graphs and implements an efficient constant factor approximation algorithm to solve it. In contrast to currently used scaffolders, it does not require either prior knowledge on the microorganisms dataset under analysis (*e.g.* their phylogenetic relationships) or the availability of paired end read libraries. This makes usability and running time two additional important features of our method. Moreover, benchmarks and tests on real bacterial datasets showed that MEDUSA is highly accurate and, in most cases, outperforms traditional scaffolders. The possibility to use MEDUSA on eukaryotic datasets has also been evaluated, leading to interesting results. [medusa/releases](#).

Genome annotation problems

Repetitive DNA, including transposable elements (TEs), is found throughout eukaryotic genomes. Annotating and assembling the “repeatome” during genome-wide analysis often poses a challenge. To address this problem, we developed DNAPIPETE – a new pipeline that uses a sample of raw genomic reads [20]. It produces precise estimates of repeated DNA content and TE consensus sequences, as well as the relative ages of TE families. We showed that DNAPIPETE performs well using very low coverage sequencing in different

genomes, losing accuracy only with old TE families. We applied this pipeline to the genome of the Asian tiger mosquito *Aedes albopictus*, an invasive species of human health interest, for which the genome size is estimated to be over 1 Gbp. Using DNAPIPETE, we showed that this species harbours a large (50% of the genome) and potentially active repeatome with an overall TE class and order composition similar to that of *Aedes aegypti*, the yellow fever mosquito. However, intra-order dynamics showed clear distinctions between the two species, with differences at the TE family level. Our pipeline's ability to manage the repeatome annotation problem will make it helpful for new or ongoing assembly projects, and our results will benefit future genomic studies of *A. albopictus*.

On another topic, we developed a reliable, robust, and much faster method for the prediction of pre-miRNAs [22]. With this method, we aimed mainly at two goals: efficiency and flexibility. Efficiency was made possible by means of a quadratic algorithm. Since the majority of the predictors use a cubic algorithm to verify the pre-miRNA hairpin structure, they may take too long when the input is large. Flexibility relies on two aspects, the input type and the organism clade. MIRINHO can receive as input both a genome sequence and small RNA sequencing (sRNA-seq) data of both animal and plant species. To change from one clade to another, it suffices to change the lengths of the stem-arms and of the terminal loop. Concerning the prediction of plant miRNAs, because their pre-miRNAs are longer, the methods for extracting the hairpin secondary structure are not as accurate as for shorter sequences. With MIRINHO, we also addressed this problem, which enabled to provide premiRNA secondary structures more similar to the ones in MIRBASE than the other available methods. MIRINHO served also as the basis to two other issues we addressed. The first issue led to the treatment and analysis of sRNA-seq data of *Acyrtosiphon pisum*, the pea aphid. The goal was to identify the miRNAs that are expressed during the four developmental stages of this species, allowing further biological conclusions concerning the regulatory system of such an organism. For this analysis, we developed a whole pipeline, called MIRINHOPIPE, at the end of which MIRINHO was aggregated. A paper is currently being prepared that presents this work.

6.3. Inferring and analysing the networks of molecular elements

Protein structure comparison

We proposed a new distance measure for comparing two protein structures based on their contact map representations [1]. We showed that our novel measure, which we refer to as the maximum contact map overlap (max-CMO) metric, satisfies all properties of a metric on the space of protein representations. Having a metric in that space allows one to avoid pairwise comparisons on the entire database and, thus, to significantly accelerate exploring the protein space compared to no-metric spaces. We showed on a gold standard superfamily classification benchmark set of 6759 proteins that our exact k -nearest neighbour ($k - NN$) scheme classifies up to 224 out of 236 queries correctly and on a larger, extended version of the benchmark with 60; 850 additional structures, up to 1361 out of 1369 queries. Our $k - NN$ classification thus provides a promising approach for the automatic classification of protein structures based on flexible contact map overlap alignments.

Metabolic network analysis

Flux balance analysis (FBA) is one of the most often applied methods on genome-scale metabolic networks. Although FBA uniquely determines the optimal yield, the pathway that achieves this is usually not unique. The analysis of the optimal-yield flux space has been an open challenge. Flux variability analysis is only capturing some properties of the flux space, while elementary mode analysis is intractable due to the enormous number of elementary modes. However, it had been previously found that the space of optimal-yield fluxes decomposes into flux modules. These decompositions allow a much easier but still comprehensive analysis of the optimal-yield flux space. Using the mathematical definition of module introduced by Müller and Bockmayr in 2013, we discovered that flux modularity is rather a local than a global property which opened connections to matroid theory [28]. Specifically, we showed that our modules correspond one-to-one to so-called separators of an appropriate matroid. Employing efficient algorithms developed in matroid theory we are now able to compute the decomposition into modules in a few seconds for genome-scale networks. Using that every module can be represented by one reaction that corresponds to its function, we also presented a method that uses this

decomposition to visualise the interplay of modules. We expect the new method to replace flux variability analysis in the pipelines for metabolic networks.

Integrated network analysis

Data on molecular interactions is increasing at a tremendous pace. Since biological functionality primarily operates at the network level, there is a clear need for topology-aware comparison methods. We developed one such method for global network alignment that is fast and robust and can flexibly deal with various scoring schemes taking both node-to-node correspondences as well as network topologies into account [18]. We exploited that network alignment is a special case of the well-studied quadratic assignment problem (QAP). We focused on sparse network alignment, where each node can be mapped only to a typically small subset of nodes in the other network. This corresponds to a QAP instance with a symmetric and sparse weight matrix. We obtained strong upper and lower bounds for the problem by improving a Lagrangian relaxation approach and introduce the open source software tool NATALIE 2.0, a publicly available implementation of our method (<https://github.com/ls-cwi/natalie>). In an extensive computational study on protein interaction networks for six different species, we found that our new method outperforms alternative established and recent state-of-the-art methods.

Integrative network analysis methods provide robust interpretations of differential high-throughput molecular profile measurements. They are often used in a biomedical context to generate novel hypotheses about the underlying cellular processes or to derive biomarkers for classification and subtyping. The underlying molecular profiles are frequently measured and validated on animal or cellular models. Therefore the results are not immediately transferable to human. In particular, this is also the case in a study of the recently discovered interleukin-17 producing helper T cells (Th17), which are fundamental for anti-microbial immunity but also known to contribute to autoimmune diseases. We proposed a mathematical model for finding active subnetwork modules that are conserved between two species [19]. These are sets of genes, one for each species, which (i) induce a connected subnetwork in a species-specific interaction network, (ii) show overall differential behaviour and (iii) contain a large number of orthologous genes. We proposed a flexible notion of conservation, which turns out to be crucial for the quality of the resulting modules in terms of biological interpretability. We developed an algorithm that finds provably optimal or near-optimal conserved active modules in our model. We applied our algorithm to understand the mechanisms underlying Th17 T cell differentiation in both mouse and human. As a main biological result, we found that the key regulation of Th17 differentiation is conserved between human and mouse.

6.4. Modelling and analysing a network of individuals, or a network of individuals' networks

Computationally investigating co-phylogenetic reconstructions and co-evolution

Despite an increasingly vast literature on co-phylogenetic reconstructions for studying host-symbiont associations, understanding the common evolutionary history of such systems remains a problem that is far from being solved. Most algorithms for host-symbiont reconciliation use an event-based model, where the events include in general (a subset of) co-speciation, duplication, loss, and host-switch. All known parsimonious event-based methods then assign a cost to each type of event in order to find a reconstruction of minimum cost. The main problem with this approach is that the cost of the events strongly influences the reconciliation obtained. To deal with this problem, we developed an algorithm, called COALA, for estimating the frequency of the events based on an approximate Bayesian computation approach [4]. The benefits of this method are twofold: (1) it provides more confidence in the set of costs to be used in a reconciliation, and (2) it allows estimation of the frequency of the events in cases where the dataset consists of trees with a large number of taxa. We evaluated our method on simulated and on biological datasets. We showed that in both cases, for the same pair of host and parasite trees, different sets of frequencies for the events lead to equally probable solutions. Moreover, often these solutions differ greatly in terms of the number of inferred events. It appears crucial to take this into account before attempting any further biological interpretation of such reconciliations. More generally, we also showed that the set of frequencies can vary widely depending on the input host and parasite trees. Indiscriminately applying a standard vector of costs may thus not be a good strategy. This work had been indicated as submitted in 2014.

Once such a cost vector has been inferred, one can proceed analysing the possible co-evolution of host-symbiont associations, phylogenetic tree reconciliation is the approach of choice for investigating the co-evolution of sets of organisms such as hosts and parasites. It consists in a mapping between the parasite tree and the host tree using event-based maximum parsimony. Given a cost model for the events, many optimal reconciliations are however possible. Only two algorithms existed that attempted such enumeration; in one case not all possible solutions are produced while in the other not all cost vectors are currently handled. We developed a polynomial-delay algorithm, EUCLYPT, for enumerating all optimal reconciliations that address these two issues [15]. We showed that in general many solutions exist. We gave an example where, for two pairs of host-parasite trees having each less than 41 leaves, the number of solutions is 5120, even when only time-feasible ones are kept. To facilitate their interpretation, those solutions are also classified in terms of how many of each event they contain. The number of different classes of solutions may thus be notably smaller than the number of solutions, yet they may remain high enough, in particular for the cases where losses have cost 0. In fact, depending on the cost vector, both numbers of solutions and of classes thereof may increase considerably (for the same instance, to respectively 4080384 and 275). To further deal with this problem, we introduced and analysed a restricted version where host-switches are allowed to happen only between species that are within some fixed distance along the host tree. This restriction allowed us to reduce the number of time-feasible solutions while preserving the same optimal cost, as well as to find time-feasible solutions with a cost close to the optimal in the cases where no time-feasible solution is found. This work had been indicated as submitted in 2014.

Evolution and metabolic complementation of organisms leaving inside the cells of another (endosymbionts)

Insect cells host many endosymbiotic bacteria, which are in general classified according to their importance for the host: “primary” symbionts are by definition mandatory and synthesise essential nutrients for the insects that feed on poor or unbalanced food sources, while “secondary” symbionts are optional and use mutualistic strategies and/or manipulation of reproduction to invade and persist within insect populations.

Hamiltonella defensa is a secondary endosymbiont that established two distinct associations with phloem-feeding insects. In aphids, it protects the host against parasitoid attacks. Its ability to infect many host tissues, notably the hemolymph, could promote its contact with parasitoid eggs. Despite this protective phenotype, the high costs associated with its presence within the host prevent its fixation in the population. In the whitefly *Bemisia tabaci* however, this symbiont is found only in cells specialised in hosting endosymbionts, the bacteriocytes. In these cells, it cohabits with other symbiotic species, such as the primary symbiont *Portiera aleyrodidarum*, a proximity that favours potential exchanges between the two symbionts. It is fixed in populations of *B. tabaci*, which suggests an important role for the consortium, probably nutritious.

We studied the specificities of each of these systems [27]. First, in the bacteriocytes of *B. tabaci*, we identified a partitioning of the synthetic capacities of two endosymbionts, *H. defensa* and *P. aleyrodidarum*, in addition to a potential metabolic complementation between the symbionts and their host for the synthesis of essential amino acids. We proposed a key nutritive role for *H. defensa*, which would indicate a transition to a mandatory status in relation to the host and would explain its fixation in the population.

We also focused on the genomic evolution of the genus *Hamiltonella*, by comparing the strains infecting *B. tabaci* with a strain infecting the aphids [29]. We highlighted the specialisation of the symbionts to their hosts, and found that the genomes of the endosymbionts reflected their respective ecology. The aphid strain thus possesses many virulence factors and is associated with two partners, a bacteriophage and a recombination plasmid. These systems, inactive in the symbiont of *B. tabaci*, are directly related to the protection against and arms race with parasitoids. Conversely, the presumed avirulence of whitefly endosymbionts is consistent with their nutritional phenotype and a transition to a mandatory status to the host.

Finally, we studied the phenomenon of “accelerated mutation rate” in *H. defensa*, compared to its sister species *Regiella insecticola*, which is also a clade of protective endosymbionts of aphids. After excluding the assumption that the transition to the intracellular life occurred independently in the two lineages, we tried to establish a link between these differences in terms of evolvability in the endosymbionts and of their gene contents, particularly for genes involved in ecology and DNA repair. All the results obtained have provided

insight into the evolution of the species *H. defensa*, since the last ancestor to the present species, by establishing a link between bacterial.

These results were part of the PhD of Pierre-Antoine Rollat-Farnier, co-supervised by Laurence Mouton (LBBE, UMR5558), Marie-France Sagot (Inria and LBBE, UMR5558) and Fabrice Vavre (LBBE, UMR5558) and defended on November 24th, 2014. The results had been indicated as submitted in 2014.

Insights on the virulence of swine respiratory tract mycoplasmas through genome-scale metabolic modelling

The respiratory tract of swines is colonised by several bacteria among which are three *Mycoplasma* species: *Mycoplasma flocculare*, *Mycoplasma hyopneumoniae* and *Mycoplasma hyorhinis*. While colonisation by *M. flocculare* was shown to be virtually asymptomatic, *M. hyopneumoniae* is known to be the causative agent of enzootic pneumonia and *M. hyorhinis* to be present in cases of pneumonia, polyserositis and arthritis. Nonetheless, the elevated genomic resemblance among these three mycoplasmas combined with their different levels of pathogenicity is an indication that they have unknown mechanisms of virulence and differential expression. We performed whole-genome metabolic network reconstructions for these three mycoplasmas and were able to show that overall they have similar metabolic capabilities. The metabolic differences that were observed include a wider range of carbohydrate uptake in *M. hyorhinis*, which in turn may also explain why this species is a widely known contaminant in cell cultures. Moreover, the myo-inositol catabolism is exclusive to *M. hyopneumoniae* and may be an important trait for virulence. However, the most important difference seems to be related to glycerol conversion to dihydroxyacetone-phosphate, which produces toxic hydrogen peroxide. This activity, missing only in *M. flocculare*, may be directly involved in cytotoxicity, as already been described for two lung pathogenic mycoplasmas, namely *Mycoplasma pneumoniae* in human and *Mycoplasma mycoides* subsp. *mycoides* in ruminants. Metabolomic data suggest that even though these mycoplasmas are extremely similar in terms of their genome and metabolism, different products and reaction rates may be the result of differential expression in each of them. We were able to infer from the reconstructed networks that the lack of pathogenicity of *M. flocculare* if compared to the highly pathogenic *M. hyopneumoniae* may be related to its incapacity to produce cytotoxic hydrogen peroxide. Moreover, the ability of *M. hyorhinis* to grow in diverse sites and even in different hosts may be a reflection of its enhanced and wider carbohydrate uptake. Altogether, the metabolic differences highlighted *in silico* and *in vitro* provide important insights to the different levels of pathogenicity observed in each of the studied species.

These results were part of the PhD of Mariana Galvão Ferrarini, co-supervised by Arnaldo Zaha (Federal University of Rio Grande do Sul and Marie-France Sagot (Inria and LBBE, UMR5558). and defended on December 10th, 2015. These results have been submitted to a journal. The PhD manuscript will be made available in HAL in early 2016.

6.5. Cross-fertilising different computational approaches

Tree matching

We considered the following problem related to tree matching, that we called the Tree-Constrained Bipartite Matching problem. Given a bipartite graph $G = (V_1, V_2, E)$ with edge weights $w : E \rightarrow \mathbb{R}^+$, a rooted tree T_1 on the set V_1 and a rooted tree T_2 on the set V_2 , find a maximum weight matching M in G , such that none of the matched nodes is an ancestor of another matched node in either of the trees [8]. This generalisation of the classical bipartite matching problem appears, for example, in the computational analysis of live cell video data. We showed that the problem is APX-hard and thus, unless $P = NP$, disproved a previous claim that it is solvable in polynomial time. Furthermore, we gave a 2-approximation algorithm based on a combination of the local ratio technique and a careful use of the structure of basic feasible solutions of a natural LP-relaxation, which we also show to have an integrality gap of $2 - o(1)$. We then considered a natural generalisation of the problem, where trees are replaced by partially ordered sets (posets). We showed that the local ratio technique gives a $2k\sigma$ -approximation for the k -dimensional matching generalisation of the problem, in which the maximum number of incomparable elements below (or above) any given element in each poset is bounded by σ . We finally gave an almost matching integrality gap example, and an inapproximability result showing that the dependence on σ is most likely unavoidable.

Graph measures

We proposed a new algorithm that computes the radius and the diameter of a weakly connected digraph $G = (V, E)$, by finding bounds through heuristics and improving them until they are validated [5]. Although the worst-case running time is $O(|V||E|)$, we experimentally showed that it performs much better in the case of real-world networks, finding the radius and diameter values after 10-100 BFSs instead of $|V|$ BFSs (independently of the value of $|V|$), and thus having running time $O(|E|)$ in practice. As far as we know, this is the first algorithm able to compute the diameter of weakly connected digraphs, apart from the naive algorithm, which runs in time $O(|V||E|)$ performing a BFS from each node. In the particular cases of strongly connected directed or connected undirected graphs, we compared our algorithm with known approaches by performing experiments on a dataset composed by several real-world networks of different kinds. These experiments showed that, despite its generality, the new algorithm outperforms all previous methods, both in the radius and in the diameter computation, both in the directed and in the undirected case, both in average running time and in robustness. Finally, as an application example, we used the new algorithm to determine the solvability over time of the “Six Degrees of Kevin Bacon” game, and of the “Six Degrees of Wikipedia” game. As a consequence, we computed for the first time the exact value of the radius and the diameter of the whole Wikipedia digraph.

The closeness and the betweenness centralities are two well-known measures of importance of a vertex within a given complex network. Having high closeness or betweenness centrality can have positive impact on the vertex itself: hence, we considered the problem of determining how much a vertex can increase its centrality by creating a limited amount of new edges incident to it [40]. We first proved that this problem does not admit a polynomial-time approximation scheme (unless $P=NP$), and we then proposed a simple greedy approximation algorithm (with an almost tight approximation ratio), whose performance is then tested on synthetic graphs and real-world networks.

The (Gromov) hyperbolicity is a topological property of a graph, which has been recently applied in several different contexts, such as the design of routing schemes, network security, computational biology, the analysis of graph algorithms, and the classification of complex networks. Computing the hyperbolicity of a graph can be very time consuming: indeed, the best available algorithm has running-time $O(n^{3.69})$, which is clearly prohibitive for big graphs. We provided a new and more efficient algorithm: although its worst-case complexity is $O(n^4)$, in practice it is much faster, allowing, for the first time, the computation of the hyperbolicity of graphs with up to 200,000 nodes [36]. We experimentally showed that the new algorithm drastically outperforms the best previously available algorithms, by analyzing a big dataset of real-world networks. Finally, we applied the new algorithm to compute the hyperbolicity of random graphs generated with the Erdős-Renyi model, the Chung-Lu model, and the Configuration Model.

Hypergraph problems

It had been previously proved independently and with different techniques that there exists an incremental output polynomial algorithm for the enumeration of the minimal edge dominating sets in graphs, *i.e.*, minimal dominating sets in line graphs. We provided the first polynomial delay and polynomial space algorithm for the problem [42]. We proposed a new technique to enlarge the applicability of Berge’s algorithm that is based on skipping hard parts of the enumeration by introducing a new search strategy. The new search strategy is given by a strong use of the structure of line graphs.

We also studied some average properties of hypergraphs and the average complexity of algorithms applied to hypergraphs under different probabilistic models [14]. Our approach is both theoretical and experimental since our goal is to obtain a random model that is able to capture the real-data complexity. Starting from a model that generalizes the Erdős-Renyi model and we obtain asymptotic estimations on the average number of transversals, irredundants and minimal transversals in a random hypergraph. We use those results to obtain an upper bound on the average complexity of algorithms to generate the minimal transversals of a hypergraph. Then we make our random model more complex in order to bring it closer to real-data and identify cases where the average number of minimal transversals is at most polynomial, quasi-polynomial or exponential.

The hypergraph transversal problem has been intensively studied, both from a theoretical and a practical point of view. In particular, its incremental complexity is known to be quasi-polynomial in general and polynomial for bounded hypergraphs. Recent applications in computational biology however require to solve a generalisation of this problem, that we call bi-objective transversal problem. The instance is in this case composed of a pair of hypergraphs (A, B) , and the aim is to enumerate minimal sets which hit all the hyperedges of A while intersecting a minimal set of hyperedges of B . We formalised this problem and related it to the enumeration of minimal hitting sets of bundles [32]. We showed cases when under degree or dimension constraints, these problems remain NP-hard, and gave a polynomial algorithm for the case when A has bounded dimension, by building a hypergraph whose transversals are exactly the hitting sets of bundles.

FLUMINANCE Project-Team

6. New Results

6.1. Fluid motion estimation

6.1.1. Stochastic uncertainty models for motion estimation

Participants: Etienne Mémin, Abed Malti.

In this study we have proposed a stochastic formulation of the brightness consistency used principally in motion estimation problems. In this formalization the image luminance is modeled as a continuous function transported by a flow known only up to some uncertainties. Stochastic calculus then enables to built conservation principles which take into account the motion uncertainties. These uncertainties defined either from isotropic or anisotropic models can be estimated jointly to the motion estimates. Such a formulation, besides providing estimates of the velocity field and of its associated uncertainties, allows us to naturally define a linear multiresolution scale-space framework. The corresponding estimator, implemented within a local least squares approach, has shown to improve significantly the results of the corresponding deterministic estimator (Lucas and Kanade estimator). This fast local motion estimator provides results that are of the same order of accuracy than state-of-the-art dense fluid flow motion estimator for particle images. The uncertainties estimated supply a useful piece of information in the context of data assimilation. This ability has been exploited to define multiscale incremental data assimilation filtering schemes. The development of an efficient GPU based version of this estimator has been investigated through the Inria ADT project FLUMILAB

6.1.2. 3D flows reconstruction from image data

Participants: Kai Berger, Cédric Herzet, Abed Malti.

Our work focuses on the design of new tools for the estimation of 3D turbulent flow motion in the experimental setup of Tomo-PIV. This task includes both the study of physically-sound models on the observations and the fluid motion, and the design of low-complexity and accurate estimation algorithms.

This year, we keep on our investigation on the problem of efficient volume reconstruction. Our work takes place within the context of some modern optimization techniques. First, we focussed our attention on the family of proximal and splitting methods and showed that the standard techniques commonly adopted in the TomoPIV literature can be seen as particular cases of such methodologies. Recasting standard methodologies in a more general framework allowed us to propose extensions of the latter: i) we showed that the parcimony characterizing the sought volume can be accounted for without increasing the complexity of the algorithms (e.g., by including simple thresholding operations); ii) we emphasized that the speed of convergence of the standard reconstruction algorithms can be improved by using Nesterov's acceleration schemes; iii) we also proposed a totally novel way of reconstructing the volume by using the so-called "alternating direction of multipliers method" (ADMM). This work has led to the publication of two contributions at the international conference on particle image velocimetry (PIV) in 2015.

On top of this work, we also focussed on another crucial step of the volume reconstruction problem, namely the pruning of the model. The pruning task consists in identifying some positions in the volume of interest which cannot contains any particle. Removing this position from the problem can then potentially allow for a dramatic dimensionality reduction. This year, we provide a methodological answer to this problem through the prism of the so-called "screening" techniques which have been proposed in the community of machine learning. Our work has led to the submission of one contribution to the international conference on acoustics, speech and signal processing.

6.1.3. Sparse-representation algorithms

Participant: Cédric Herzet.

The paradigm of sparse representations is a rather new concept which turns out to be central in many domains of signal processing. In particular, in the field of fluid motion estimation, sparse representation appears to be potentially useful at several levels: i) it provides a relevant model for the characterization of the velocity field in some scenarios; ii) it plays a crucial role in the recovery of volumes of particles in the 3D Tomo-PIV problem.

Unfortunately, the standard sparse representation problem is known to be NP hard. Therefore, heuristic procedures have to be devised to access to the solution of this problem. Among the popular methods available in the literature, one can mention orthogonal matching pursuit (OMP), orthogonal least squares (OLS) and the family of procedures based on the minimization of ℓ_p norms. In order to assess and improve the performance of these algorithms, theoretical works have been undertaken in order to understand under which conditions these procedures can succeed in recovering the "true" sparse vector.

This year, we contributed to this research axis by deriving conditions of success for the algorithms mentioned above when the amplitudes of the nonzero coefficients in the sparse vector obey some decay. In a TomoPIV context, this decay corresponds to the fact that not all the particles in the fluid diffuse the same quantity of light (notably because of illumination or radius variation). In particular, we show that the standard coherence-based guarantees for OMP/OLS can be relaxed by an amount which depends on the decay of the nonzero coefficients. Our work have led to the acceptance of a paper in the journal IEEE Transactions on Information Theory.

6.2. Tracking, Data assimilation and model-data coupling

6.2.1. Sequential smoothing for fluid motion

Participants: Anne Cuzol, Etienne Mémin.

In parallel to the construction of stochastic filtering techniques for fluid motions, we have proposed a new sequential smoothing method within a Monte-Carlo framework. This smoothing aims at reducing the temporal discontinuities induced by the sequential assimilation of discrete time data into continuous time dynamical models. The time step between observations can indeed be long in environmental applications for instance, and much longer than the time step used to discretize the model equations. While the filtering aims at estimating the state of the system at observations times in an optimal way, the objective of the smoothing is to improve the estimation of the hidden state between observation times. The method is based on a Monte-Carlo approximation of the filtering and smoothing distributions, and relies on a simulation technique of conditional diffusions. The proposed smoother can be applied to general non linear and multidimensional models. It has been applied to a turbulent flow in a high-dimensional context, in order to smooth the filtering results obtained from a particle filter with a proposal density built from an Ensemble Kalman procedure. This conditional simulation framework can also be used for filtering problem with low measurement noise. This has been explored through a collaboration with Jean-Louis Marchand (ENS Bretagne) in the context of vorticity tracking from image data.

6.2.2. Stochastic fluid flow dynamics under uncertainty

Participants: Etienne Mémin, Valentin Resseguier.

In this research axis we aim at devising Eulerian expressions for the description of fluid flow evolution laws under uncertainties. Such an uncertainty is modeled through the introduction of a random term that allows taking into account large-scale approximations or truncation effects performed within the dynamics analytical constitution steps. This includes for instance the modeling of unresolved scales interaction in large eddies simulation (LES) or in Reynolds average numerical simulation (RANS), but also uncertainties attached to non-uniform grid discretization. This model is mainly based on a stochastic version of the Reynolds transport theorem. Within this framework various simple expressions of the drift component can be exhibited for different models of the random field carrying the uncertainties we have on the flow. We aim at using such a formalization within image-based data assimilation framework and to derive appropriate stochastic versions of geophysical flow dynamical modeling. This formalization has been published in the journal Geophysical

and Astrophysical Fluid Dynamics [9]. Numerical simulation on divergence free wavelets basis of 3D viscous Taylor-Green vortex and Crow instability have been performed within a collaboration with Souleymane Kadri-Harouna. Besides, we explore in the context of Valentin Resseguier's PhD the extension of such framework to oceanic models and to satellite image data assimilation. This PhD thesis takes place within a fruitful collaboration with Bertrand Chapron (CERSAT/IFREMER). This year we have more deeply explored several uncertainty representations of classical geophysical models for ocean and atmosphere. This study have led to very promising stochastic representation for the Quasi Geostrophic approximation (QG) with noises of different energy.

6.2.3. *Free surface flows reconstruction and tracking*

Participants: Dominique Heitz, Etienne Mémin.

We investigated the combined use of a Kinect depth sensor and of a stochastic data assimilation method to recover free-surface flows. More generally, we proposed a particle filter method to reconstruct the complete state of free-surface flows from a sequence of depth images only. The data assimilation scheme introduced accounts for model and observations errors. We evaluated the developed approach on two numerical test cases: a collapse of a water column as a toy-example and a flow in an suddenly expanding flume as a more realistic flow. The robustness of the method to simulated depth data quality and also to initial conditions was considered. We illustrated the interest of using two observations instead of one observation into the correction step. Then, the performance of the Kinect sensor to capture temporal sequences of depth observations was investigated. Finally, the efficiency of the algorithm was qualified for a wave in a real rectangular flat bottom tank. It was shown that for basic initial conditions, the particle filter rapidly and remarkably reconstructed velocity and height of the free surface flow based on noisy measures of the elevation

6.2.4. *Optimal control techniques for the coupling of large scale dynamical systems and image data*

Participants: Pranav Chandramouli, Dominique Heitz, Etienne Mémin, Cordelia Robinson.

In this axis of work we are exploring the use of optimal control techniques for the coupling of Large Eddies Simulation (LES) techniques and 2D image data. The objective is to reconstruct a 3D flow from a set of simultaneous time resolved 2D image sequences visualizing the flow on a set of 2D plans enlightened with laser sheets. This approach will be experimented on shear layer flows and on wake flows generated on the wind tunnel of Irstea Rennes. Within this study we wish also to explore techniques to enrich large-scale dynamical models by the introduction of uncertainty terms or through the definition of subgrid models from the image data. This research theme is related to the issue of turbulence characterization from image sequences. Instead of predefined turbulence models, we aim here at tuning from the data the value of coefficients involved in traditional LES subgrid models or in longer-term goal to learn empirical subgrid models directly from image data. An accurate modeling of this term is essential for Large Eddies Simulation as it models all the non resolved motion scales and their interactions with the large scales.

We have pursued the first investigations on a 4DVar assimilation technique, integrating PIV data and Direct Numerical Simulation (DNS), to reconstruct two-dimensional turbulent flows. The problem we are dealing with consists in recovering a flow obeying Navier-Stokes equations, given some noisy and possibly incomplete PIV measurements of the flow. By modifying the initial and inflow conditions of the system, the proposed method reconstructs the flow on the basis of a DNS model and noisy measurements. The technique has been evaluated in the wake of a circular cylinder. It denoises the measurements and increases the spatiotemporal resolution of PIV time series. These results have been recently published in the Journal of Computational Physics [6]. Along the same line of studies the 3D case is ongoing. The goal consists here to reconstruct a 3D flow from a set of simultaneous time resolved 2D images of planar sections of the 3D volume. This work has been mainly conducted within the PhD of Cordelia Robinson. The development of the variational assimilation code has been initiated within a collaboration with A. Gronskis, S. Laizé (lecturer, Imperial College, UK) and Eric Lamballais (institut P' Poitiers). A High Reynolds number simulation of the wake behind a cylinder has been recently performed within this collaboration. The 4DVar assimilation technique based on the numerical code Incompact3D is now implemented. We are currently trying to reconstruct a 3D turbulent flow from dual

plane velocity observations. The control of subgrid parameterizations will be the main objective of the PhD of Pranav Chandramouli that is just starting.

6.2.5. Ensemble variational data assimilation of large scale fluid flow dynamics with uncertainty

Participant: Etienne Mémin.

This study is focused on the coupling of a large scale representation of the flow dynamics built from the location uncertainty principle with image data of finer resolution. The velocity field at large scales is described as a regular smooth component whereas the complement component is a highly oscillating random velocity field defined on the image grid but living at all the scales. Following this route we have assessed the performance of an ensemble variational assimilation technique with direct image data observation. Preliminary encouraging results have been obtained for simulation under uncertainty of 1D and 2D shallow water models.

6.2.6. Reduced-order models for flows representation from image data

Participants: Cédric Herzet, Etienne Mémin, Valentin Resseguier.

During the PhD thesis of Valentin Resseguier we proposed a new decomposition of the fluid velocity in terms of a large-scale continuous component with respect to time and a small-scale non continuous random component. Within this general framework, an uncertainty based representation of the Reynolds transport theorem and Navier-Stokes equations can be derived, based on physical conservation laws. This physically relevant stochastic model has been applied in the context of the POD-Galerkin method. The pertinence of this reduced order model has been successfully assessed on several wake flows. This study has been published in two conference papers and one journal article.

On the other hand, we also investigated the problem of reduced-model construction from partial observations. In this line of search, our contribution was twofold. We first proposed a Bayesian framework for the construction of reduced-order models from image data. Our framework enables to account for any prior information on the system to reduce and takes the uncertainties on the parameters of the model into account. Interestingly, the proposed approach reduces to some well-known model-reduction techniques when the observations are not partial (i.e., the observation operator can be inverted). Second, we provided a theoretical analysis of our methodology in a simplified context (namely, the observations are supposed to be noiseless linear combinations of the state of the system). This result provides worst-case guarantees on the reconstruction performance which can be achieved by a reduced model built from the data. These contributions have been accepted for presentation in two international conferences in 2016.

6.3. Analysis and modeling of turbulent flows

6.3.1. Turbulence similarity theory for the modeling of Ocean Atmosphere interface

Participants: Roger Lewandowski, Etienne Mémin, Benoit Pinier.

The Ocean Atmosphere interface plays a major role in climate dynamics. This interaction takes place in a thin turbulent layer. To date no satisfying universal models for the coupling of atmospheric and oceanic models exists. In practice this coupling is realized through empirically derived interaction bulks. In this study, corresponding to the PhD thesis of Benoit Pinier, we aim at exploring similarity theory to identify universal mean profile of velocity and temperature within the mixture layer. The goal of this work consists in exhibiting eddy viscosity models within the primitive equations. We will also explore the links between those eddy viscosity models and the subgrid tensor derived from the uncertainty framework studied in the Fluminance group. In that prospect, we have started to study the impact of the introduction of a random modeling of the friction velocity on the classical wall law expression.

6.3.2. Hot-wire anemometry at low velocities

Participant: Dominique Heitz.

A new dynamical calibration technique has been developed for hot-wire probes. The technique permits, in a short time range, the combined calibration of velocity, temperature and direction calibration of single and multiple hot-wire probes. The calibration and measurements uncertainties were modeled, simulated and controlled, in order to reduce their estimated values. Based on a market study the french patent application has been extended this year to a Patent Cooperation Treaty (PCT) application.

6.3.3. Numerical and experimental image and flow database

Participants: Pranav Chandramouli, Dominique Heitz.

The goal was to design a database for the evaluation of the different techniques developed in the Fluminance group. The first challenge was to enlarge a database mainly based on two-dimensional flows, with three-dimensional turbulent flows. Synthetic image sequences based on homogeneous isotropic turbulence and on circular cylinder wake have been provided. These images have been completed with time resolved Particle Image Velocimetry measurements in wake and mixing layers flows. This database provides different realistic conditions to analyse the performance of the methods: time steps between images, level of noise, Reynolds number, large-scale images. The second challenge was to carried out orthogonal dual plane time resolved stereoscopic PIV measurements in turbulent flows. The diagnostic employed two orthogonal and synchronized stereoscopic PIV measurements to provide the three velocity components in planes perpendicular and parallel to the streamwise flow direction. These temporally resolved planar slices observations will be used in 4DVar assimilation technique, integrating Direct Numerical Simulation (DNS) and Large Eddies Simulation (LES), to reconstruct three-dimensional turbulent flows. This reconstruction will be conducted within the PhD of Pranav Chandramouli. The third challenge was to carried out a time resolved tomoPIV experiments in a turbulent wake flow. These temporally resolved volumic observations will be used to assess the algorithms developped in the PhD of Ioana Barbu and in the postdoc of Kai Berger. Then this data will be used in 4DVar assimilation technique to reconstruct three-dimensional turbulent flows. This reconstruction will be conducted within the PhD of Cordelia Robinson.

6.4. Visual servoing approach for fluid flow control

6.4.1. Closed-loop control of a spatially developing shear layer

Participant: Christophe Collewet.

This study is led within a strong collaboration with Diemer Ando-Ondo and Johan Carlier of the Acta team (Irstea Rennes). It aims at controlling one of the prototypical flow configurations encountered in fluid mechanics: the spatially developing turbulent shear layer occurring between two parallel incident streams with different velocities. Closed loop control is achieved to maintain the shear-layer in a desired state of interest for industrial applications, and thus to reject upstream perturbations. The industrial and scientific contexts advocates first for the use of image sensor to measure the flow velocity fields and second for applying the control on the upstream boundary condition. The optimal control was performed using a linear control law designed from a reduced linearized state space model of the Navier-Stokes equations. A steady desired state was first considered leading to a linear time-invariant system. The resulting feedback control law was validated on a powerful and realistic numerical Navier-Stokes 3D solver, which will be useful to anticipate the control of the shear layer in a dedicated wind tunnel. Two conference papers on this work have been submitted to the "16th European Control Conference" and "8th AIAA Flow Control Conference". We are now considering the case of an unsteady desired state to control the large roller vortices developing in the shear layer and that are the main contributor to entrainment and mixing processes.

GALEN Project-Team

7. New Results

7.1. Optimizing Average Precision

Participants: Pawan Kumar

Average precision (AP) is one of the most commonly used measures for ranking. However, due to the inefficiency of optimizing it during learning, a common approach is to use surrogate loss functions such as 0-1 loss. We have developed a novel latent AP-SVM classifier [1], that minimizes a carefully designed upper bound on the AP-based loss function over weakly supervised samples. Using publicly available datasets, we demonstrate the advantage of our approach over standard loss-based learning frameworks on three challenging problems: action classification, character recognition and object detection.

7.2. Region-based Semantic Segmentation

Participants: Pawan Kumar

In [9] we consider the problem of parameter estimation and energy minimization for a region-based semantic segmentation model. The main problem we face in the context of energy minimization, is the large number of putative pixel-to-region assignments. We address this problem by designing an accurate linear programming based approach for selecting the best set of regions from a large dictionary, which is constructed by merging and intersecting segments obtained from multiple bottom-up over-segmentations. The lack of fully supervised data is tackled by using a latent structural SVM formulation, where the latent variables model any missing information in the human annotation. Using large, publicly available datasets we show that our methods are able to significantly improve the accuracy of the region-based model.

7.3. Parsimonious Labeling

Participants: Puneet Dokania, Pawan Kumar

In [22], we propose a new family of discrete energy minimization problems, which we call parsimonious labeling, that is to use as few labels as possible. This allows us to capture useful cues for important computer vision applications such as stereo correspondence and image denoising. Furthermore, we propose an efficient graph-cuts based algorithm for the parsimonious labeling problem that provides strong theoretical guarantees on the quality of the solution. Using both synthetic and standard real datasets, we show that our algorithm significantly outperforms other graph-cuts based approaches.

7.4. Structured Learning and Inference

Participants: Jiaqian Yu, Matthew Blaschko

We have developed computationally efficient structured output prediction methods for learning with non-modular losses [19], [29], [40]. We both demonstrate the feasibility of learning with submodular losses, as well as show that learning with multiple correct outputs can lead to NP-hard optimization problems even when learning with a single correct output is feasible.

7.5. Novel graph kernels

Participants: Katerina Gkirtzou, Matthew Blaschko

We have developed a novel family of graph kernels that are capable of incorporating local curvature properties of 3D meshes [6]. We have additionally demonstrated their application to the modelling of interdependencies between different brain regions in an fMRI based classification task for predicting cocaine addiction.

7.6. Structured Sparsity Regularization & Statistical Hypothesis Testing

Participants: Eugene Belilovsky, Wacha Bounliphone, Katerina Gkirtzou, Andreas Argyriou, Matthew Blaschko

We have developed novel methods for structured sparsity regularization & hypothesis testing. We have applied these methods to fMRI [3], [2], [36] and the analysis of large medical databases [10]. We have also developed novel statistical hypothesis tests for relative dependency [21], [37] and similarity [14]. We have applied these methods to the problem of identifying relative dependencies between languages using a multi-lingual corpus, and for discovering the relative relationships between gliomas and genetic information. Additionally, we have shown the application of relative tests to the problem of model selection in deep generative models, and currently an important question in machine learning.

7.7. High-Order MRF models

Participants: Nikos Komodakis, Nikos Paragios

We developed a very general algorithm for structured prediction learning [7] that is able to efficiently handle discrete MRFs/CRFs (including both pairwise and higher-order models) so long as they can admit a decomposition into tractable subproblems. By properly combining dual decomposition with a max-margin learning method, the framework manages to reduce the training of a complex high-order MRF to the parallel training of a series of simple slave MRFs that are much easier to handle.

7.8. Graph-based registration and segmentation

Participants: Enzo Ferrante, Vivien Fecamp, Aimilia Gastouniotti, Bharat Singh, Stavros Alchatzidis, Nikos Paragios

Deformable image registration plays a fundamental role in many clinical applications. We investigated the use of graphical models in the context of a particular type of image registration problem, known as slice-to-volume registration. We introduced a scalable, modular and flexible formulation that can accommodate low-rank [5] and high order [16] terms, that simultaneously selects the plane and estimates the in-plane deformation through a single shot optimization approach. We applied our models on simulated and real-data in the context of ultrasound and magnetic resonance registration, demonstrating the potential of our methods.

We also developed a novel methodology for graph-based motion-driven segmentation [24] and applied it for carotid plaque segmentation in ultrasound images. We identified the plaque region by exploiting kinematic dependencies between the atherosclerotic and the normal arterial wall. The methodology exploits group-wise image registration towards recovering the deformation field, on which information theory criteria are used to determine dominant motion classes and a map reflecting kinematic dependencies, which is then segmented using Markov random fields.

Moreover, in order to address the problem of general purpose multi-modal deformable registration/fusion we developed a novel and robust method using a metric defined in an appropriate sub-space which is adaptive to the image-content/image-modality [18]. We adopted a graph-based formulation that assumes that intensities of corresponding pixels in the two image domains are related through an unknown piece-wise constant linear function. This relation is propagated to an appropriate sub-space (wavelets coefficients) where a criterion that couples the estimation of the deformation field with optimal transport function on the subspace and the smoothness of the deformation is considered.

7.9. Object Detection from RGB-Depth images

Participant: Siddhartha Chandra, Iasonas Kokkinos

In [11] we explore RGB-Depth representations for the training of Deformable Models, and describe strategies to improve an object detection pipeline by introducing viewpoint based mixture components. Our contributions are threefold. First, we use surface-based object representations (3D mesh models) from available 3D object model repositories to learn strongly supervised viewpoint classifiers. Second, we develop a geometric dataset augmentation scheme that uses scene geometry to ‘take another look’ at the training data, simulating the effect of camera viewpoint changes. Third, to better exploit depth information, we develop a novel depth-based dense feature extraction method that provides a robust statistical description of scene geometry. We evaluate our learned detectors on the common NYU dataset, and demonstrate that each of our advances results in systematic performance improvements over the traditional detection pipeline.

7.10. Deep CNN for Modelling Deformations and Semantic Segmentation

Participant: Iasonas Kokkinos

Invariance to deformations in Deep Convolutional Neural Networks (DCNN) is commonly achieved by using multiple ‘max-pooling’ (MP) layers. In [26] we show that alternative methods of modeling deformations can improve the accuracy and efficiency of DCNNs. For this, (i) we introduce epitomic convolution as an alternative to the common convolution-MP cascade of DCNNs, (ii) we introduce a Multiple Instance Learning algorithm to accommodate global translation and scaling in image classification and (iii) we develop a DCNN sliding window detector that explicitly, but efficiently, searches over the object’s position, scale, and aspect ratio. We provide competitive image classification and localization results on the ImageNet dataset and object detection results on Pascal VOC2007.

In [25] we bring together methods from DCNNs and probabilistic graphical models for addressing the task of pixel-level classification (“semantic image segmentation”). We overcome the poor localization property of deep networks by combining the responses at the final DCNN layer with a fully connected Conditional Random Field (CRF). Qualitatively, our “DeepLab” system is able to localize segment boundaries at a level of accuracy which is beyond previous methods.

7.11. Learning Low-level Image Representations with Deep CNN

Participant: Iasonas Kokkinos

In [27] we propose a novel framework for learning local image descriptors in a discriminative manner. For this purpose we explore a siamese architecture of DCNNs, with a Hinge embedding loss on the L2 distance between descriptors. Since a siamese architecture uses pairs rather than single image patches to train, there exist a large number of positive samples and an exponential number of negative samples. We propose to explore this space with a stochastic sampling of the training set, in combination with an aggressive mining strategy over both the positive and negative samples which we denote as “fracking”. We perform a thorough evaluation of the architecture hyper-parameters, and demonstrate large performance gains compared to both standard CNN learning strategies, hand-crafted image descriptors like SIFT, and the state-of-the-art on learned descriptors: up to 2.5x vs SIFT and 1.5x vs the state-of-the-art in terms of the area under the curve (AUC) of the Precision-Recall curve.

In [4] we explore connections between DCNNs and texture understanding. First, instead of focusing on texture instance and material category recognition, we propose a human-interpretable vocabulary of texture attributes to describe common texture patterns, complemented by a new describable texture dataset for benchmarking. Second, we look at the problem of recognizing materials and texture attributes in realistic imaging conditions, including when textures appear in clutter, developing corresponding benchmarks on top of the recently proposed OpenSurfaces dataset. Third, we revisit classic texture representations, including bag-of-visual-words and the Fisher vectors, in the context of deep learning and show that these have excellent efficiency and generalization properties if the convolutional layers of a deep model are used as filter banks. We obtain in this manner state-of-the-art performance in numerous datasets well beyond textures, an efficient method to apply deep features to image regions, as well as benefit in transferring features from one domain to another.

In [35] we propose a new DCNN architecture that learns pixel embeddings, such that pairwise distances between the embeddings can be used to infer whether or not the pixels lie on the same region. That is, for any two pixels on the same object, the embeddings are trained to be similar; for any pair that straddles an object boundary, the embeddings are trained to be dissimilar. Experimental results show that when this embedding network is used in conjunction with a DCNN trained on semantic segmentation, there is a systematic improvement in per-pixel classification accuracy. Our contributions are integrated in the popular Caffe deep learning framework, and consist in straightforward modifications to convolution routines. As such, they can be exploited for any task involving convolution layers.

7.12. Human-Limb Segmentation for Intelligent Mobility Assistance Robots

Participants: Siddhartha Chandra, Stavros Tsogkas, Iasonas Kokkinos

We developed a computer vision component [12] to be used as part of an intelligent robotic assistant. This component exploits RGB and depth information extracted from Kinect sensors mounted on the robot, to accurately segment human limbs, using fully convolutional neural networks (FCNNs). We trained our network using an in-house Human-Limb dataset composed of video frames, and described a scheme for dynamically exploiting RGB and depth data in a single framework for training and testing. Our method demonstrated promising performance, being very efficient at the same time, with a run-time of 8 frames per second on a single GPU.

GENSCALE Project-Team

7. New Results

7.1. HTS data processing

7.1.1. Genome Analysis Tool Box Optimization

Participants: C. Deltel, P. Durand, E. Drezen, D. Lavenier, C. Lemaitre, P. Peterlongo, G. Rizk

Among the GATB library, the kmer-counting procedure is one of the most useful building block to speed-up development of new NGS tools. It is the first step of many NGS tools developed with GATB : Leon, Bloocoo, MindTheGap, DiscoSnp, Simka, TakeAbreak. This procedure has been optimized to be less limited by disks I/O. It relies on the use of kmer minimizers that help quickly partition the whole set of kmers into compact subsets. The kmer-counting procedure has also been re-worked to be more versatile, it is now able to count separately many input files and allows easy parametrization of the output, from simple kmer-count to the creation of custom user-defined kmer measures. At the core of the GATB library is also the manipulation and traversal of the de Bruijn Graph. The implementation has been optimized, leading to graph traversal twice fast as before. We introduced a new type of bloom filters, that are specially optimized for the manipulation of kmers. In these bloom filters neighboring kmers in the graph are close together in the bloom filter bit array, leading to better data locality, less cache misses and better overall performance [38].

7.1.2. NGS Data Compression

Participants: G. Benoit, E. Drezen, D. Lavenier, C. Lemaitre, G. Rizk

A novel reference-free method to compress data issued from high throughput sequencing technologies has been developed. Our approach, implemented in the LEON software, employs techniques derived from assembly principles. The method is based on a reference probabilistic de-Bruijn Graph, built de novo from the set of reads and stored in a Bloom filter. Each read is encoded as a path in this graph, by memorizing an anchoring kmer and a list of bifurcations. The same probabilistic de-Bruijn Graph is used to perform a lossy transformation of the quality scores allowing higher compression rates to be obtained without losing pertinent information for downstream analyses. Leon was run on various real sequencing datasets (whole genome, exome, RNA-seq or metagenomics). In all cases, LEON showed higher overall compression ratios than state-of-the-art compression software. On a *C. elegans* whole genome sequencing dataset, LEON divided the original file size by more than 20 [16].

7.1.3. Multistep global optimization approach for the scaffolding problem

Participants: R. Andonov, D. Lavenier, I. Petrov

Our overall goal here is to address the computational hardness of the scaffolding problem by designing faster algorithms for global optimization that combine the branch-and-bound method which is able to find the global optimum but is usually slow for accuracy, with the use of massive parallelism and exploiting the special properties of the data—for scalability. A new two step scaffolding modeling strategy is in development. It tries to break the problem complexity by first solving a graph containing only large unitigs building something that can be compared to a trustworthy genomic frame. In our preliminary works [40] we developed integer programming optimization models that have been successfully applied on synthetic data generated from small chloroplast genomes. For computation we use the Gurobi optimization solver.

7.1.4. Mapping reads on graph

Participants: A. Limasset, C. Lemaitre, P. Peterlongo

Next Generation Sequencing (NGS) has dramatically enhanced our ability to sequence genomes, but not to assemble them. In practice, many published genome sequences remain in the state of a large set of contigs. Although many subsequent analyses can be performed, one may ask whether mapping reads on the contigs is as informative as mapping them on the paths of the assembly graph. We proposed a formal definition of mapping on a de Bruijn graph, analysed the problem complexity which turned out to be NP-complete, and provided a practical solution. We proposed a pipeline called GGMAP (Greedy Graph MAPping). Its novelty is a procedure to map reads on branching paths of the graph, for which we designed a heuristic algorithm called BGREAT (de Bruijn Graph READ mapping Tool). For the sake of efficiency, BGREAT rewrites a read sequence as a succession of unitigs sequences. GGMAP can map millions of reads per CPU hour on a de Bruijn graph built from a large set of human genomic reads. Surprisingly, results show that up to 22% more reads can be mapped on the graph but not on the contig set. Although mapping reads on a de Bruijn graph is a complex task, our proposal offers a practical solution combining efficiency with an improved mapping capacity compared to assembly-based mapping even for complex eukaryotic data [43].

7.1.5. Improving discoSnp features

Participants: C. Riou, C. Lemaitre, P. Peterlongo

NGS data enable to detect polymorphisms such as SNPs and indels. Their detection in NGS data is now a routine task. The main methods for their prediction usually need a reference genome. However, non-model organisms and highly divergent genomes such as in cancer studies are more and more investigated. The discoSnp tool has been successfully applied to predict isolated SNPs from raw read set(s) without the need of a reference genome. We improved discoSnp which became discoSnp++ [44]. DiscoSnp++ benefits from a new software design that reduces time and memory consumption, and from a new algorithmic design that detects all kinds of SNP and small indels, adds genotype information and outputs a VCF (Variant Calling Format) file. Moreover, when a reference genome may be used, discoSnp++ predictions are automatically mapped to this reference and the VCF file shows up location information of each prediction. This step also provides a way to filter out false predictions due to genomic repeats. Using discoSnp++ even when a reference is available has multiple advantages: it is several order of magnitude faster and uses much less memory. We are currently working in showing that it also provides better predictions than methods based on read mapping.

7.1.6. HLA genotyping

Participant: D. Lavenier

The human leukocyte antigen (HLA) system drives the regulation of the Human immune system. Genotyping the HLA genes involved in the immune system consists first in a deep sequencing of the HLA region. Next, a NGS analysis is performed to detect SNP variations from which correct haplotypes are computed. We have developed a fast method that outperforms standard approaches which, generally, require exhaustive database searches. Instead, the method extracts a few significant k-mers from all the haplotypes referenced in the HLA database. Each haplotype is then characterized by a small set of informative k-mers. By comparing these k-mer sets with the HLA sequencing data of a specific person, we can rapidly determine its HLA genotype.

7.1.7. Identification of long non-coding RNAs in insects genomes

Participant: F. Legeai

The development of high throughput sequencing technologies (HTS) has allowed researchers to better assess the complexity and diversity of the transcriptome. Among the many classes of non-coding RNAs (ncRNAs) that were identified during the last decade, long non-coding RNAs (lncRNAs) represent a diverse and numerous repertoire of important ncRNAs, reinforcing the view that they are of central importance to the cell machinery in all branches of life. Although lncRNAs have been involved in essential biological processes such as imprinting, gene regulation or dosage compensation especially in mammals, the repertoire of lncRNAs is poorly characterized for many non-model organisms [23]. In collaboration with the Institut de Génétique et de Développement de Rennes (IGDR) we participate in the development of a software for extracting long non coding RNA from high throughput data (<https://github.com/tderrien/FEELnc>).

7.1.8. Data-mining applied to GWAS

Participants D. Lavenier, Pham Hoang Son

Discriminative pattern mining methods are powerful techniques for discovering variant combinations related to diseases. The aim is to find a set of patterns that occur with disproportion frequency in case-control data sets, and a real challenge is to select a complete set of variant combinations that are biologically significant. There are various measurement methods for evaluating the discriminative power of individual combination in two-class data sets. Our research activity on this topic attempts to compare the statistical discriminative power measurements in genetic case-control data sets in order to evaluate the effectiveness of detecting variants associated with diseases.

7.2. Sequence comparison

7.2.1. Amplicon alignment

Participants: S. Brillet, C. Deltel, P. Durand, D. Lavenier, I. Petrov

Many metagenomics projects identify species by the studying 16S-RNA sequences. This is mainly done by comparing the amplicons with 16S-RNA bacterial banks (amplicons are short fragments sequenced from very specific genome areas). As these sequences share a lot of similarities, immediate blast-like heuristics achieve poor performances. To speed up the process, we first select informative k-mers, from both the amplicon dataset and in the RNA16 bank (informative k-mers are defined as under represented k-mers). An index is built from this reduced set of k-mers and a "seed-and-extend" procedure is run. This strategy avoids many non-useful computation and accelerate the overall computation by two orders of magnitude. This new approach is currently implemented in the PLAST software (Regional KoriPlast2 project).

7.2.2. Metagenomics datasets comparison

Participants: G. Benoit, D. Lavenier, C. Lemaitre, P. Peterlongo, G. Rizk

We develop a new method, called Simka, to compare simultaneously numerous large metagenomics datasets. The method computes pairwise distances based on the amount of shared k-mers between datasets. The method scales to a large number of datasets thanks to an efficient kmer-counting step that processes all datasets simultaneously. Additionally, several distance definitions were implemented and compared, including some originating from the ecological domain. The method is currently applied to the TARA oceans project (more than 500 datasets) which aims at comparing worldwide sea water samples (ANR HydrGen project) [39].

7.3. Protein 3D structure

7.3.1. Discovering protein conformations by distance geometry

Participant: A. Mucherino

The distance geometry asks whether a simple weighted undirected graph G can be embedded in a Euclidean space having a predefined dimension $K > 0$, so that distances between pairs of embedded vertices are the same as the weights on graph edges. One of the most important applications of the distance geometry can be found in biology, where experimental techniques are able to find estimates of certain distances between atom pairs in molecules. Even if the scientific community is used to employ standardized techniques for the solution of this problem, which are essentially based on heuristic searches, we have recently shown that our combinatorial approach to this problem can be in fact employed for solving biological instances of the distance geometry [17]. This work is in collaboration with international people and researchers from the Pasteur Institut in Paris.

7.3.2. Discretization orders for distance geometry

Participant: A. Mucherino

The concept of discretization order is fundamental for the discretization of the distance geometry, i.e. for reducing the search space of a given distance geometry instance to a discrete (and finite) space. A discretization order is an order on the vertices of the graph G representing an instance of the distance geometry that is able to satisfy the discretization assumptions. Recent research was focused on the problem of finding, for a given distance geometry instance, a suitable discretization order that allows for its discretization [32]. The problem is tackled from a purely theoretical point of view in [33], while a special order for protein backbones was identified in [27] by creating a path on a "pseudo" de Bruijn graph. In [36], additional requirements are included during the search for a vertex order, in order to identify discretization orders that are also "optimal". In this work, we used Answer Set Programming (ASP) for identifying optimal partial orders that ensure the discretization of distance geometry instances related to proteins. This work is in collaboration with the Dyliss team, as well as international people.

7.3.3. Structure Similarity Detection

Participants: M. Le Boudic-jamin, R. Andonov

The most commonly used among the various measures of alignment similarity are the internal distances root mean squared deviation (RMSDd) and the coordinate root mean squared deviation (RMSDc). In the paper [18] we introduce a novel approach to find similarities between protein structures. Our algorithm is both internal-distances based and Euclidean-coordinates based (i.e., it uses a rigid transformation to optimally superimpose the two structures). Resulting alignments are guaranteed to score well for both RMSDd and RMSDc, while remaining polynomial. We also replace the goal of finding the largest clique by the one of returning several very dense "near-clique" subgraphs. This choice is strongly justified by the observation that distinct solutions to the structural alignment problem that are close to the optimum are all equally viable from the biological perspective, and hence are all equally interesting from the computation standpoint. Our tool is suitable for detecting similar domains when comparing multi-domain proteins, as well to detect structural repetitions within a single protein and between related proteins [12].

7.3.4. Automatic Classification of Protein Structure

Participants: M. Le Boudic-jamin, R. Andonov

In this paper [15] we propose a new distance measure for comparing two protein structures based on their contact map representations. We show that our novel measure, which we refer to as the maximum contact map overlap (max-CMO) metric, satisfies all properties of a metric on the space of protein representations. Having a metric in that space allows one to avoid pairwise comparisons on the entire database and, thus, to significantly accelerate exploring the protein space compared to no-metric spaces. We show on a gold standard superfamily classification benchmark set of 6759 proteins that our exact k-nearest neighbor (k-NN) scheme classifies up to 224 out of 236 queries correctly and on a larger, extended version of the benchmark with 60 850 additional structures, up to 1361 out of 1369 queries. Our k-NN classification thus provides a promising approach for the automatic classification of protein structures based on flexible contact map overlap alignments.

7.3.5. Detection of structure repeats in proteins

Participant: M. Le Boudic-jamin, R. Andonov

Almost 25% of proteins contain internal repeats, these repeats may have a major role in the protein function. Furthermore some proteins actually are the same substructure repeated many times, these proteins are solenoids. However, very few protein repeats detection programs exist today. In the paper [29] we present a simple and efficient tool for discovering protein repeats. Our tool is based on protein fragment comparison and clique detection. We show that our tool is able to detect different levels of repetitions and to successfully identify protein tiles.

7.4. Parallelism

7.4.1. Processor in Memory

Participants: C. Deltel, D. Lavenier

The concept of PIM (Processor In Memory) aims to dispatch the computer power near the data. Together with the UPMEM company, which is currently developing a DRAM enhanced with computing units, we investigate the parallelization of several bioinformatics algorithms for this new types of memory. The first results show that blast-like algorithms or mapping algorithms can highly benefit of such memory. But the core algorithms must be revisited in order to better suite the PIM architecture.

7.4.2. Alignment search tools on cloud

Participants: S. Brillet, D. Lavenier, I. Petrov

PLAST is an alternative version of Blast to target intensive sequence comparison (bank-to-bank comparison). The multicore version offers a speed from 5 to 10 compared to Blast. In 2015, we deploy PLAST in the IFB cloud infra-structure (French Bioinformatics Institute) and demonstrate that an Hadoop implementation provides a very good scalability [34].

7.4.3. Bioinformatics Workflow

Participants: D. Lavenier, F. Moorews

Bioinformatics workflows play an important role in the development of new methodologies for analyzing sequencing data. Optimizing this activity brings the questions of how workflow can be efficiently captured and how technical tasks integration can be simplified. Thus, we define an expressive graphic workflow language, adapted to the quick capture of workflows. This graphical input is then interpreted by a workflow engine based on a new model of computation with high performances obtained by the use of multiple levels of parallelism. A Model-Driven design approach is associated to facilitate the data parallelism generation and the production of suitable implementations for different execution contexts. In the case of the cloud model Container as a Service (CaaS), a workflow specification intrinsically re-executable and readily disseminatable has been developed. The adoption of this kind of model could lead to an acceleration of exchanges and a better availability of data analysis workflows [25] [31] [13].

7.4.4. Graph processing

Participants: D. Lavenier, R. Andonov

In the paper [20] we present a new approach for solving the all-pairs shortest-path (APSP) problem for planar graphs that exploits the massive on-chip parallelism available in today's Graphics Processing Units (GPUs). We describe two new algorithms based on our approach. Both algorithms use Floyd-Warshall method, have near optimal complexity in terms of the total number of operations, while their matrix-based structure is regular enough to allow for efficient parallel implementation on the GPUs. By applying a divide-and-conquer approach, we are able to make use of multi-node GPU clusters, resulting in more than an order of magnitude speedup over fastest known Dijkstra-based GPU implementation and a two-fold speedup over a parallel Dijkstra-based CPU implementation.

7.4.5. Analytical models and Optimization for GPUs

Participants: R. Andonov

In [28] we develop a methodology for modeling the energy efficiency of tiled nested-loop codes running on a graphics processing unit (GPU) and use it for energy efficiency optimization. We use the polyhedral model, and we assume that a highly optimized and parametrized version of a tiled nested – loop code, either written by an expert programmer or automatically produced by a polyhedral compilation tool – is given to us as an input. We then model the energy consumption as an analytical function of a set of parameters characterizing the software and the GPU hardware. Our approach develops analytical models based on (i) machine and architecture parameters, (ii) program size parameters as found in the polyhedral model and (iii) tiling parameters, such as those that are chosen by auto-or manual tuners. Our model therefore allows efficient optimization of the energy efficiency with respect to a set of parameters of interest.

7.5. Applications

7.5.1. CAMI: Critical Assessment of Metagenomic Interpretation

Participants: C. Deltel, D. Lavenier, C. Lemaitre, P. Peterlongo, G. Rizk

The interpretation of metagenomes relies on sophisticated computational approaches such as short read assembly, binning and taxonomic classification. All subsequent analyses can only be as meaningful as the outcome of these initial data processing methods⁰. The CAMI initiative aims to evaluate these methods independently, comprehensively and without bias. The goal is to supply users with exhaustive quantitative data about the performance of methods in many relevant scenarios. In 2015, we participate to CAMI within the "assembly" category using the Minia assembly pipeline. Results are provided here: <https://data.cami-challenge.org/>. For the medium challenge datasets, our assemblies are referred under the identifiers *goofy-wilson* and *fervent-blackwell*.

7.5.2. Assembly and Annotation of Arthropods Genomes

Participants: A. Gouin, F. Legeai, C. Lemaitre

Within a large international network of biologists, GenScale has contributed to various projects for identifying important components such as protein coding or non coding genes involved in the adaptation of major agricultural pests to their environment. We provided the assembly and the annotation of 4 new aphids, 3 parasitic wasps, and improved the assembly of 2 variants of fall army worm by removing unwanted sequences due to heterozygosity [41], [42]. Following specific agreement or policy, these new genomes and annotations are available for a restricted consortium or a large community through the BioInformatics platform for Agro-ecosystems Arthropods (<http://bipaa.genouest.org/is>). These results, and further analyses led to a better understanding of the biology, evolution and life history traits of *Spodoptera frugiperda* [19], the identification and characterization of new genome of pea aphid symbionts [22] and the identification of differentially expressed genes in the sensory system of *Sesamia nonagrioides* [21].

7.5.3. Study of the rapeseed genome structure

Participants: D. Lavenier, C. Lemaitre, S. Letort, P. Peterlongo

In collaboration with IGEPP (Institut de Génétique, Environnement et Protection des Plantes), INRA, and through two national projects, PIA Rapsodyn and France-Génomique Polysuccess, we are involved in the genome analysis of several rapeseed varieties. The Rapsodyn project has the ambition to insure long-term competitiveness of the rapeseed production through improvement of the oil yield and reduction of nitrogen inputs during the crop cycle. Rapeseed varieties must thus be selected from genotypes that favor low nitrogen input. DiscoSNP++ is here used to locate new variants among the large panel of rapeseed varieties which have been sequenced during the project. The PolySuccess project aims to answer the following question: how a polyploid, such as the oilseed rape plant, becomes a new species? Oilseed rape (*Brassica napus*) being a natural hybrid between *B.rapa* and *B.oleracea*, different genomes of these three species have been sequenced to study their structures. The Minia assembly pipeline provides a fast way to generate contigs that are used for studying gene specificities.

⁰<http://www.cami-challenge.org/>

IBIS Project-Team

6. New Results

6.1. Inference of bacterial regulatory networks from reporter gene data

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.2). In order to fully exploit this technology, we need methods to rapidly construct reporter genes, both on plasmids and on the chromosome, mathematical models to infer biologically relevant quantities from the primary data, and computer tools to achieve this in an efficient and user-friendly manner. For instance, in a typical microplate experiment, 96 cultures are followed in parallel, over several hours, resulting in 10,000-100,000 measurements of absorbance and fluorescence and luminescence intensities. Over the past few years, we put into place an experimental platform and data analysis software, notably the WELLREADER program (Section 5.4), to allow biologists to make the most out of the information contained in reporter gene expression data. An invited review on the analysis of fluorescent reporter gene data was published in the proceedings of the Third International Workshop on Hybrid Systems Biology (HSB 14) [25].

Valentin Zulkower, in the framework of his PhD thesis, has developed novel methods for the analysis of reporter gene data, based on the use of regularized linear inversion. This allows a range of estimation problems in the analysis of reporter gene data, notably the inference of growth rate, promoter activity, and protein concentration profiles, to be solved in a mathematically sound and practical manner. We have evaluated the validity of the approach using *in-silico* simulation studies, and observed that the methods are more robust and less biased than indirect approaches usually encountered in the experimental literature based on smoothing and subsequent processing of the primary data, like in WELLREADER. We have applied the methods to the analysis of fluorescent reporter gene data acquired in kinetic experiments with *Escherichia coli*. The methods were shown capable of reliably reconstructing time-course profiles of growth rate, promoter activity, and protein concentration from weak and noisy signals at low population volumes. Moreover, they captured critical features of those profiles, notably rapid changes in gene expression during growth transitions. The linear inversion methods have been implemented in the Python package WELLFARE, and integrated by Michel Page in the web application WELLINVERTER (Section 5.3). This work was presented at the major bioinformatics conference ISMB/ECCB 2015 and published in the special issue of *Bioinformatics* associated with the conference [24]. The Institut Français de Bioinformatique (IFB) accepted a proposal to extend WellInverter into a scalable and user-friendly web service providing a guaranteed quality of service, in terms of availability and response time. This web service will be deployed on the IFB platform and accompanied by extensive user documentation, online help, and a tutorial.

Over the years, the above tools have been used in several studies in IBIS directed at the experimental mapping of gene regulatory networks in *E. coli*. An example is the motility network of *E. coli*, studied by Diana Stefan in the context of her PhD thesis. The main thrust of this work lies in clarifying and solving methodological issues in the automated inference of quantitative models of gene regulatory networks from time-series gene expression data, also called reverse engineering in the bioinformatics literature. The application of existing reverse engineering methods is commonly based on implicit assumptions on the biological processes under study. First, the measurements of mRNA abundance obtained in transcriptomics experiments are taken to be representative of protein concentrations. Second, the observed changes in gene expression are assumed to be solely due to transcription factors and other specific regulators, while changes in the activity of the gene expression machinery and other global physiological effects are neglected. While convenient in practice, these assumptions are often not valid and bias the reverse engineering process. In her PhD thesis, Diana Stefan systematically investigated, using a combination of models and experiments, the importance of this bias and possible corrections. She measured with the help of fluorescent reporter genes the activity of genes involved in the FliA-FlgM module of the *E. coli* motility network. From these data, protein concentrations and global physiological effects were estimated by means of kinetic models of gene expression. The results indicate

that correcting for the bias of commonly-made assumptions improves the quality of the models inferred from the data. Moreover, it was shown by simulation that these improvements are expected to be even stronger for systems in which protein concentrations have longer half-lives and the activity of the gene expression machinery varies more strongly across conditions than in the FliA-FlgM module. The approach proposed in this study is broadly applicable when using time-series transcriptome data to learn about the structure and dynamics of regulatory networks. The paper describing the work was published in *PLoS Computational Biology* [23].

In addition to reporter gene data, a variety of other experimental data can be used for the mapping of gene regulatory networks. For example, using Chromatin Immunoprecipitation-sequencing (ChIP-seq) experiments, Stéphane Lacour and colleagues have identified a large number of target promoters of the sigma factor σ^S during the transition from exponential to stationary phase. Sigma factors are accessory subunits of RNA polymerase, allowing the recognition of specific promoter sequences by the transcriptional machinery, and σ^S is known to specifically accumulate in a variety of stress conditions. The study, published in *Scientific Reports* [21], has confirmed the importance of σ^S for redirecting RNA polymerase to promoters that drive the expression of genes necessary for the survival of *E. coli* after nutrient exhaustion. Furthermore, the results highlight the role of σ^S in the regulation of several noncoding RNAs.

6.2. Models of carbon metabolism in bacteria

All free-living bacteria have to adapt to a changing environment. Specific regulatory systems respond to particular stresses, but the most common decision bacteria have to make is the choice between alternative carbon sources, each sustaining a specific, maximal growth rate. Many bacteria have evolved a strategy that consists in utilizing carbon sources sequentially, in general favouring carbon sources that sustain a higher growth rate. As long as a preferred carbon source is present in sufficient amounts, the synthesis of enzymes necessary for the uptake and metabolism of less favourable carbon sources is repressed. This phenomenon is called Carbon Catabolite Repression (CCR) and the most salient manifestation of this regulatory choice is diauxic growth, a phenomenon discovered by Jacques Monod more than 70 years ago. Although this system is one of the paradigms of the regulation of gene expression in bacteria, the underlying mechanisms remain controversial. CCR involves the coordination of different subsystems of the cell - responsible for the uptake of carbon sources, their breakdown for the production of energy and precursors, and the conversion of the latter to biomass.

The complexity of this integrated system, with regulatory mechanisms cutting across metabolism, gene expression, signaling and subject to global physical and physiological constraints, has motivated important modeling efforts over the past four decades, especially in the enterobacterium *Escherichia coli*. Different hypotheses concerning the dynamic functioning of the system have been explored by a variety of modeling approaches. In an article in *Trends in Microbiology* [19], which was initiated during the sabbatical of Andreas Kremling in Grenoble in 2013, we have reviewed these studies and summarized their contributions to the quantitative understanding of CCR, focusing on diauxic growth in *E. coli*. Moreover, we have proposed a highly simplified representation of diauxic growth that makes it possible to bring out the salient features of the models proposed in the literature and confront and compare the explanations they provide. In parallel, specific aspects of CCR, in particular a better understanding of the role of the signalling molecule cyclic adenosine monophosphate (cAMP) in the dynamic regulation of promoters during growth transitions in *E. coli*, have been studied in the context of the PhD thesis of Valentin Zulkower, using both models and experimental data.

Beside CCR and the multiple regulatory systems controlling the metabolism of *E. coli*, the involvement of post-transcriptional regulation is uncertain. The post-transcriptional factor CsrA is stated as being the only regulator essential for the use of glycolytic substrates, but its impact on the functioning of central carbon metabolism has not been demonstrated. In the framework of the PhD thesis of Manon Morin, supported by a Contrat Jeune Scientifique INRA-Inria, the collaboration of Delphine Ropers, Muriel Coccagn-Bousquet and Brice Enjalbert from LISBP at INSA Toulouse has resulted in a multi-scale analysis of a wild-type strain and its isogenic mutant attenuated for CsrA. A variety of experimental data has been acquired for these two strains in relevant conditions, including growth parameters, gene expression levels, metabolite pools, enzyme

activities and metabolic fluxes. Data integration, metabolic flux analysis and regulation analysis revealed the pivotal role of post-transcriptional regulation for reshaping carbon metabolism. In particular, the work has shed light on *csrA* essentiality and has provided an explanation for the glucose-phosphate stress observed in the mutant strain. A paper summarizing the work has been submitted for publication in a microbiology journal.

6.3. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single-cell level, the processes that govern single-cell dynamics in general and gene expression in particular are better described by stochastic models. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling ranges from a better comprehension of the biochemical regulatory mechanisms underlying cellular phenotypes to the development of new strategies for the (computer assisted or genetically engineered) control of cell populations and even of single cells.

Work in IBIS on gene expression and interaction dynamics at the level of individual cells is addressed in terms of identification of intrinsic noise models from population snapshot data, on the one hand, and the inference of models focusing on cellular variability within isogenic populations from fluorescence microscopy gene expression profiles, on the other hand. Along with modelling and inference comes analysis of the inferred models in various respects, notably in terms of identifiability, single-cell state estimation and control. Other problems related with single-cell modelling and extracellular variability are considered in eukaryotic cells through external collaborations.

In the context of the response of yeast cells to osmotic shocks, in collaboration with the LIFEWARE project-team and colleagues from Université Paris Descartes and University of Pavia (Italy), Eugenio Cinquemani has investigated the use of mixed effects-modelling and identification techniques to characterize individual cell dynamics in isogenic cell populations. Mixed-effects models are hierarchical models where parametric response profiles of individuals is subject to inter-individual parameter variability following a common population distribution. Starting from identification approaches in pharmacokinetics, we have developed and applied inference methods to microfluidics data, with a focus on the response of budding yeast to osmotic shocks. First results presented at conference in 2013 and the identification and validation work performed with Andres Gonzales, who visited IBIS for a few months in 2014 during his PhD at the University of Pavia, have been finalized into a journal article recently accepted for publication in *PLoS Computational Biology* [20].

Started with a study of the arabinose uptake dynamics in *E. coli*, work on identification and state estimation for single-cell intrinsic noise models of gene networks has focused on the reconstruction of promoter activity profiles from fluorescent reporter data. In the single-cell stochastic context, given population snapshots of fluorescence levels at subsequent experimental instants, the problem becomes that of inferring promoter activity statistics over a cell population such as mean, variance or even higher-order moments from analogous statistics of the reporter output. This nontrivial extension of the deterministic deconvolution of promoter activity from population-average data requires knowledge of the stochastic reporter dynamics and of the relation between promoter and fluorescence statistical moments. In two conference papers, Eugenio Cinquemani investigated identifiability and identification of the kinetic parameters of the stochastic reporter dynamics [28] and proposed parametric and nonparametric methods for the reconstruction of the desired promoter activity statistics [27], [28], demonstrating their effectiveness *in silico*. Further developments of these methods and application to experimental data for addressing relevant biological questions will be the subject of future journal publications.

In parallel, collaboration of Eugenio Cinquemani with Marianna Rapsomaniki, post-doctoral researcher at at IBM Zurich Research Lab (Switzerland), Zoi Lygerou at the University of Patras (Greece) and John Lygeros at ETH Zurich (Switzerland) has been devoted to the analysis of data from Fluorescence Recovery After Photobleaching (FRAP) experiments and the inference of kinetic parameters of protein dynamics in single

eukaryotic cells. As an alternative to current approximate analytical methods, we have explored inference methods based on simulation of biological processes in realistic environments at a particle level. We introduced and demonstrated a new method for the inference of kinetic parameters of protein dynamics, where a limited number of *in-silico* FRAP experiments is used to construct a mapping from FRAP recovery curves to the parameters sought. Parameter estimates from experimental data are then computed by applying the mapping to the observed recovery curves, at virtually no additional price for any number of experiments, along with the application of a bootstrap procedure for determining identifiability of the parameters and confidence intervals for their estimates. After validation on synthetic data, the method was successfully applied to the analysis of the nuclear proteins Cdt1, PCNA and GFPnls in mammalian cells, also shedding light on cell-to-cell variability of the protein kinetics. Method and results have been published in *Bioinformatics* this year [22].

6.4. Growth control in bacteria and biotechnological applications

The ability to experimentally control the growth rate is crucial for studying bacterial physiology. It is also of central importance for applications in biotechnology, where often the goal is to limit or even arrest growth. Growth-arrested cells with a functional metabolism open the possibility to channel resources into the production of a desired metabolite, instead of wasting nutrients on biomass production. The objective of the RESET project, supported in the framework of the Programme d'Investissements d'Avenir (Section 8.2), is to develop novel strategies to limit or completely stop microbial growth and to explore biotechnological applications of these approaches.

A foundation result for growth control in bacteria was published in the journal *Molecular Systems Biology* this year [18]. In this publication, which is based on the PhD thesis of Jérôme Izard and post-doctoral work of Cindy Gomez Balderas, we describe an engineered *E. coli* strain where the transcription of a key component of the gene expression machinery, RNA polymerase, is under the control of an inducible promoter. By changing the inducer concentration in the medium, we can adjust the RNA polymerase concentration and thereby switch bacterial growth between zero and the maximal growth rate supported by the medium. We have shown that our synthetic growth switch functions in a medium-independent and reversible way, and we have provided evidence that the switching phenotype arises from the ultrasensitive response of the growth rate to the concentration of RNA polymerase. In parallel, Delphine Ropers in collaboration with Jean-Luc Gouzé and Stefano Casagrande of the BIOCORE team are developing a quantitative model of the gene expression machinery to account for this surprising observation.

The publication in *Molecular Systems Biology* also presents a biotechnological application of the growth switch in which both the wild-type *E. coli* strain and our modified strain are endowed with the capacity to produce glycerol when growing on glucose. Cells in which growth has been switched off continue to be metabolically active and harness the energy gain to produce glycerol at a twofold higher yield than in cells with natural control of RNA polymerase expression. Remarkably, without any further optimization, the improved yield is close to the theoretical maximum computed from a flux balance model of *E. coli* metabolism. The synthetic growth switch is thus a promising tool for gaining a better understanding of bacterial physiology and for applications in synthetic biology and biotechnology. We submitted a patent for such applications at the European Patent Office.

Whereas the synthetic growth switch has been designed for biotechnological purposes, the question can be asked how resource allocation is organized in wild-type strains that have naturally evolved. Recent work has shown that coarse-grained models of resource allocation can account for a number of empirical regularities relating the macromolecular composition of the cell to the growth rate. Some of these models hypothesize control strategies enabling microorganisms to optimize growth. While these studies focus on steady-state growth, such conditions are rarely found in natural habitats, where microorganisms are continually challenged by environmental fluctuations. The aim of the PhD thesis of Nils Giordano is to extend the study of microbial growth strategies to dynamical environments, using a self-replicator model. In a recently submitted paper, we have formulated dynamical growth maximization as an optimal control problem that can be solved using Pontryagin's Maximum Principle. We compare this theoretical gold standard with different possible

implementations of growth control in bacterial cells. This study has been carried out in collaboration with Jean-Luc Gouzé and Francis Mairet of the BIOCORE project-team.

LEMON Team

7. New Results

7.1. Hydrodynamics of the Tunquen lagoon, Chile

Participant: Antoine Rousseau.

In this internship co-advised with Céline Acary-Robert (Inria Chile), Loïc Dagnas developed a numerical hydrodynamic model for a specific lagoon of the Chilean coastline. This kind of lagoon is characterized by an intermittent connection to the sea and a regular fresh water input coming from the Andean mountains. The hydrodynamic model consists in a two-dimensional shallow water model, including tracer equations for the time evolution of temperature and salinity. The hydrodynamic circulation of the lagoon has been simulated taking into account various external forcings such as water exchanges with the atmosphere, wind effects and external pumping.

7.2. Upscaled modeling of Vaccares lake in Camargue

Participants: Carole Delenne, Antoine Rousseau, Vincent Guinot.

Sélim Cornet developed a numerical model for the hydrodynamics of Vaccares system in Camargue. The data and reference simulations (made with TELEMAC-2D) were provided by Tour du Valat (contact O. Boutron). Sélim's work consisted in the implementation and validation of the porosity shallow water model developed by Vincent GUINOT, in order to obtain accurate but inexpensive simulations of the Vaccares hydrosystem.

7.3. Numerical simulation of coastal flood made by Joanna storm (France, 2008)

Participant: Fabien Marche.

In collaboration with BRGM, a numerical platform based on Fabien Marche's numerical tool WaveBox was developed in order to simulate coastal urban submersions associated with intense storms, see [22]. A nudging strategy is implemented with:

- a barotropic model at the regional scale,
- a spectral wave model with embedded meshes accounting for water level evolution and output from the large scale model,
- a free surface Shallow Water model (SURF2D, now called WaveBox) used at very high resolution for the submersion process.

7.4. Analysis of the inclusion of vorticity on fully nonlinear and weakly dispersive long wave models

Participant: Fabien Marche.

We study in [11] the propagation of long waves in the presence of vorticity. In the irrotational framework, the Green-Naghdi equations (also called Serre or fully nonlinear Boussinesq equations) are the standard model for the propagation of such waves. These equations couple the surface elevation to the vertically averaged horizontal velocity and are therefore independent of the vertical variable. In the presence of vorticity, the dependence on the vertical variable cannot be removed from the vorticity equation but it was however shown in [9] that the motion of the waves could be described using an extended Green-Naghdi system. In this paper we propose an analysis of these equations, and show that they can be used to get some new insight into wave-current interactions. We show in particular that solitary waves may have a drastically different behavior in the presence of vorticity and show the existence of solitary waves of maximal amplitude with a peak at their crest, whose angle depends on the vorticity. We also propose a robust and simple numerical scheme validated on several examples. Finally, we give some examples of wave-current interactions with a non trivial vorticity field and topography effects.

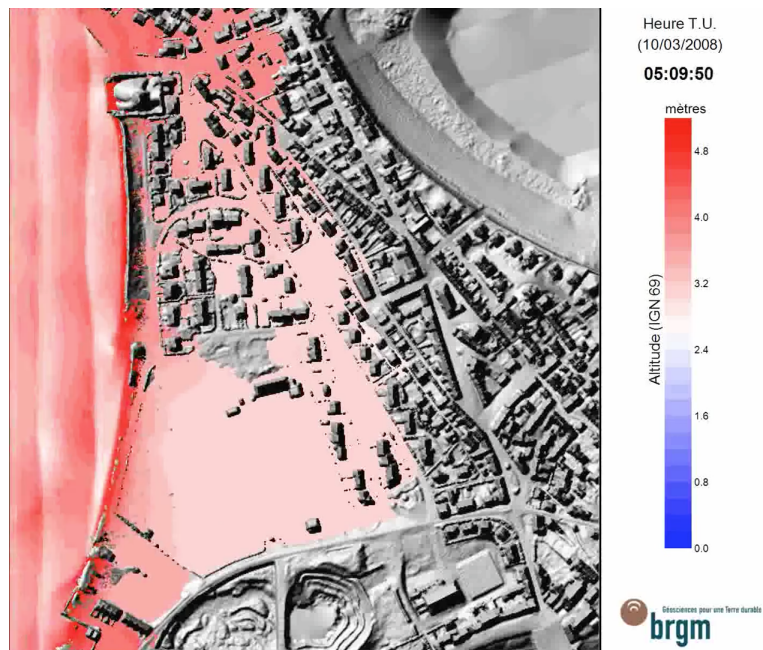


Figure 4. Simulation of Joanna storm (France, 2008). Snapshot of the movie made by BRGM, available on their [Youtube page](#).

7.5. Multiscale aspects for confinement of coastal lagoons

Participant: Antoine Rousseau.

In [3] we expand a previous definition of paralic confinement (see [41]), and make it usable from the modeling slant, before implementing it in numerical tools. More specifically, we here deal with the multiscale aspect of the confinement. If a paralic environment is separated into two (or more) connected areas, we will show that it is possible to split the confinement problem into two related problems, one for each area. We also focus on the importance of the interface length between the two subdomains.

7.6. Interface conditions for ocean models

Participant: Antoine Rousseau.

In [5] we are interested in the search of interface conditions to couple hydrostatic and nonhydrostatic ocean models. To this aim, we consider simplified systems and use a time discretization to handle linear equations. We recall the links between the two models (with the particular role of the aspect ratio $\delta = H/L$) and introduce an iterative method based on the Schwarz algorithm (widely used in domain decomposition methods). The convergence of this method depends strongly on the choice of interface conditions: this is why we look for exact absorbing conditions and their approximations in order to provide tractable and efficient coupling algorithms.

In [4] we present a study of optimized Schwarz domain decomposition methods for Navier-Stokes equations. Once discretized in time, optimal transparent boundary conditions are derived for the resulting Stokes equations, and a series of local approximations for these nonlocal conditions are proposed. Their convergence properties are studied, and numerical simulations are conducted on the test case of the driven cavity. It is shown that conditions involving one or two degrees of freedom can improve the convergence properties of the original algorithm.

7.7. Use of remote sensing data for hydraulic modelling

Participant: Carole Delenne.

Wetlands provide a vital resource to ecosystem services and associated rural livelihoods; but their extent, geomorphological heterogeneity and flat topography make the representation of their hydrological functioning complex. The main objective of this area of research is to assess the relevance of remote sensing data for the monitoring and hydraulic modelling of different hydrosystems. In [14], a semi automated method exploiting 526 MODIS 8-day 500 m resolution images was developed to study the spatial and temporal dynamics of the annual flood across the Niger Inner Delta over the period 2000-2011. The flooded area is detected using band ratio indexes. Results were evaluated against classified Landsat images, previous studies and field stage data for a range of hydrological units: river stretches, lakes, floodplains and irrigated areas. Depending on the study area, its extent, and the objective to be reached, different kinds of remote sensing data may be interesting: RADAR, multispectral, high/low spatial/temporal resolution, etc. Several paths for research are currently considered to upgrade the use of remote-sensing data in hydrodynamic modelling:

- use of detected flooded area for model validation and for the calibration of parameters such as friction coefficient.
- topography assessment using the detection of the flooded area of a given wetland at different times.
- characterization of statistical properties of the geometry of the urban medium (useful for large-scale models): statistical, subgrid-scale properties of the topography, and information regarding the flow connectivity properties of the urban medium.

7.8. Lumped hydrological models with infinite characteristic time transfer functions

Participant: Vincent Guinot.

Karst and mountainous catchments usually exhibit rainfall-runoff transfer functions involving multiple time scales. In most existing conceptual, hydrological models of such catchments, multiple time scale response is achieved by introducing several reservoirs and non-linear transfer functions. In [9], multiple time scales are introduced by proposing a transfer function with an infinite characteristic time. The heavy-tailed transfer function behaves asymptotically as an inverse power of time. In the limit of long time scales, the governing equation for the system obeys a fractional differential equation. With a single reservoir, the proposed approach is shown to perform satisfactorily compared to other models of similar or more complex structure. The fractional differential equation is shown to be useless for usual time scales and should not be used in practice.

7.9. Upscaled models for urban floods

Participant: Vincent Guinot.

Shallow water models with porosity have arisen over the last two decades as a promising alternative to refined flow models for the simulation of urban floods. Several porosity-based models have been proposed in the literature. In [10], the integral porosity formalism developed at the University of California Irvine is validated against scale model experiments. A sudden dike breaching near an idealized city layout is simulated in the scale model. Comparison with numerical simulations shows the superiority of the integral porosity model over the single porosity model in reproducing the effects of urban layout anisotropy on flood wave propagation properties. This research has initiated a collaboration between the LEMON team and UC Irvine for the development of a new porosity formalism.

7.10. Models for dispersion in porous media

Participants: Carole Delenne, Vincent Guinot.

Solute dispersion in porous media is usually modelled using Fick's law or fractional variations of the solute dispersion equation. The Fickian model, however, is known to exhibit a number of drawbacks, such as poor scaling properties. This is also true for its fractional counterparts, that perform with limited success when compared to experimental data sets. In [13], a high-quality experimental device is built in the form of periodic heterogeneities of length 15 cm. Placing up to 10 periods in series allows the scaling properties of the dispersion model to be analyzed. Besides providing a high quality experimental database, the results in [13] indicate that (i) previously identified scaling trends for the dispersion coefficient may easily be explained by experiment variability, (ii) there exists a linear transport model that allows the experimental behaviour to be reproduced at all scales, (iii) this model is not the advection-dispersion model (even fractional).

7.11. Invasion in growth-fragmentation-death models

Participant: Fabien Campillo.

In collaboration with Nicolas Champagnat and Coralie Fritsch (Inria Nancy), we present in [20] two approaches to study invasion in growth-fragmentation-death models: one based on a stochastic individual based model and one based on an integro-differential model. The invasion of the population is described by the survival probability for the first model and by an eigenproblem for the second one. We study these two notions of invasion fitness, giving different characterizations of the growth of the population, and we make links between these two complementary points of view. We apply our work in the context of adaptive dynamics in a chemostat model.

7.12. Stochastic growth model with extinction

Participant: Fabien Campillo.

In collaboration with Marc Joannides and Irène Larramendy-Valverde (IMAG / Université de Montpellier), we consider in [6] a stochastic logistic growth model given by a stochastic differential equation featuring both birth and death rates in the drift and diffusion coefficients. Our aim is to infer these rates, based on discrete observations with possible extinction. Since extinction occurs eventually for the model, the density of the diffusion process is not absolutely continuous with respect to the Lebesgue measure; we established the associated Fokker-Planck equation together with appropriate numerical schemes. This formulation allows to design variants of the standard methods that can handle extinction.

LIFEWARE Project-Team

7. New Results

7.1. Hybrid Simulation of Heterogeneous Biochemical Models in SBML

Participants: Katherine Chiang, François Fages, Sylvain Soliman.

Models of biochemical systems presented as a set of formal reaction rules can be interpreted in different formalisms, most notably as either deterministic Ordinary Differential Equations, stochastic continuous-time Markov Chains, Petri nets or Boolean transition systems. While the formal composition of reaction systems can be syntactically defined as the (multiset) union of the reactions, the composition and simulation of models in different formalisms remains a largely open issue. In [5], we show that the combination of reaction rules and events, as already present in SBML, can be used in a non-standard way to define stochastic and boolean simulators and give meaning to the hybrid composition and simulation of heterogeneous models of biochemical processes. In particular, we show how two SBML reaction models can be composed into one hybrid continuous-stochastic SBML model through a high-level interface for composing reaction models and specifying their interpretation. Furthermore, we describe dynamic strategies for automatically partitioning reactions with stochastic or continuous interpretations according to dynamic criteria. The performances are then compared to static partitioning. The proposed approach is illustrated and evaluated on several examples, including the reconstructions of the hybrid model of the mammalian cell cycle regulation of Singhanian et al. as the composition of a Boolean model of cell cycle phase transitions with a continuous model of cyclin activation, the hybrid stochastic-continuous models of bacteriophage T7 infection of Alfonsi et al., and the bacteriophage λ model of Goutsias, showing the gain in both accuracy and simulation time of the dynamic partitioning strategy.

7.2. Theoretical and Practical Complexities of Enumerating Minimal Siphons in Petri Nets

Participants: François Fages, Thierry Martinez, Sylvain Soliman.

Petri nets are a simple formalism for modeling concurrent computation. They are also an interesting tool for modeling and analysing biochemical reaction systems, bridging the gap between purely qualitative and quantitative models. Biological networks can indeed be complex, large, and with many unknown kinetic parameters, which makes the development of quantitative models difficult. In [9], we focus on the Petri net representation of biochemical reactions and on two structural properties of Petri nets, siphons and traps, that bring us information about the persistence of some molecular species, independently of the kinetics. We first study the theoretical time complexity of minimal siphon decision problems in general Petri nets, and present three new complexity results: first, we show that the existence of a siphon of a given cardinality is NP-complete; second, we prove that deciding the Siphon-Trap property is co-NP-complete; third, we prove that deciding the existence of a minimal siphon containing a given set of places, deciding the existence of a siphon of a given cardinality and deciding the Siphon-Trap property can be done in linear time in Petri nets of bounded tree-width. Then, we present a Boolean model of siphons and traps, and two methods for enumerating all minimal siphons and traps of a Petri net, by using a SAT solver and a Constraint Logic Program (CLP) respectively. On a benchmark of 345 Petri nets of hundreds of places and transitions, extracted from biological models from the BioModels repository, as well as on a benchmark composed of 80 Petri nets from the Petriweb database of industrial processes, we show that both the SAT and CLP methods are overall faster by one or two orders of magnitude compared to the state-of-the-art algorithm from the Petri net community, and are in fact able to solve all the enumeration problems of our practical benchmarks. We investigate why these programs perform so well in practice, and provide some elements of explanation related to our theoretical complexity results.

7.3. Abstraction-based Parameter Synthesis for Multiaffine Systems

Participant: Grégory Batt.

Multiaffine hybrid automata (MHA) represent a powerful formalism to model complex dynamical systems. This formalism is particularly suited for the representation of biological systems which often exhibit highly non-linear behavior. In [10], we consider the problem of parameter identification for MHA. We present an abstraction of MHA based on linear hybrid automata, which can be analyzed by the SpaceEx model checker. This abstraction enables a precise handling of time-dependent properties. We demonstrate the potential of our approach on a model of a genetic regulatory network and a myocyte model.

7.4. Tropical Algebra Methods for Model Reduction

Participants: François Fages, Jonas Sénizergues, Sylvain Soliman.

Jonas Sénizergues has just started a PhD Thesis on the design of model reduction techniques for systems biology based on tropical algebra. The idea is to reason on the orders of magnitude of both kinetic parameters and molecular concentrations in order to determine particular regimes exhibiting fast-slow decomposition and amenable to model reductions. Such model reductions generalize the quasi steady-state (QSSA) and quasi-equilibrium (QE) criteria, and lead to hybrid automata for chaining the reduced dynamics. The solving of tropical equilibration equations rely on previous work using constraint programming techniques⁰ with collaboration with Ovidier Radulescu (Univ. Montpellier) and Andreas Weber (University of Bonn, Germany).

7.5. Modeling the Effect of the Cell Cycle on the Circadian Clock in Mouse Embryonic Fibroblasts

Participants: François Fages, Jonas Sénizergues, Denis Thieffry, Pauline Traynard, Sylvain Soliman.

Experimental observations have put in evidence autonomous self-sustained circadian oscillators in most mammalian cells, and proved the existence of molecular links between the circadian clock and the cell cycle. Several models have been elaborated to assess conditions of control of the cell cycle by the circadian clock, in particular through the regulation by clock genes of Wee1, an inhibitor of the mitosis promoting factor, responsible for a circadian gating of mitosis and cell division period doubling phenomena. However, recent studies in individual NIH3T3 fibroblasts have shown an unexpected acceleration of the circadian clock together with the cell cycle when the milieu is enriched in FBS, the absence of such acceleration in confluent cells, and the absence of any period doubling phenomena. In [14], we try to explain these observations by a possible entrainment of the circadian clock by the cell cycle through the inhibition of transcription during mitosis. We develop a differential model of that reverse coupling of the cell cycle and the circadian clock and investigate the conditions in which both cycles are mutually entrained. We use the mammalian circadian clock model of Relogio et al. and a simple model of the cell cycle by Qu et al. which focuses on the mitosis phase. We show that our coupled model is able to reproduce the main observations reported by Feillet et al. in individual fibroblast experiments and use it for making some predictions. In [17], those hypothesis are revised in order to reproduce the phase data in addition to the period data and make new predictions.

7.6. Effects of repeated osmotic stress on gene expression and growth: from cell-to-cell variability to cellular individuality in the budding yeast *Saccharomyces cerevisiae*

Participants: Grégory Batt, Ewen Corre, Pascal Hersen, Artémis Llamosi.

⁰Sylvain Soliman, François Fages, Ovidiu Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014.

When shifted to a stressful environment, cells are capable of complex response and adaptations. Although the cellular response to a single stress has been studied in great detail, very little is known when it comes to dynamically fluctuating stressful environments. In addition, in the context of stress response, the role of cell-to-cell variability in cellular processes and more specifically in gene expression is still unclear.

In his PhD thesis [3], Art emis Llamosi uses a systems and synthetic biology approach to investigate osmotic stress in *S. cerevisiae* at the single cell level. Combining microfluidics, fluorescent microscopy and advanced image analysis, we are able to subject cells to precise fluctuating osmolarity and monitor single-cell temporal response.

While much previous research in gene expression heterogeneity focused on its stochastic aspect, we consider here long-lasting differences between cells regarding expression kinetics. Using population models and state-of-the-art statistical analysis, we manage to represent both population and single-cell dynamics in a single concise modelling framework. This quantitative approach capturing stable individuality in gene expression dynamics can define a form of non-genetic cellular identity.

To improve our understanding of the biological interpretation of such identity, we investigate the relation between single-cell specificities in their gene expression with their phenotype and micro-environment. We then take a lineage based perspective and find this form of identity to be partially inherited.

Understanding the evolutionary consequences of inheritable non-genetic cellular identity requires a better knowledge of the impact of fluctuating stress on cell proliferation. Dissecting quantitatively the consequences of repeated stress on cell-cycle and growth gives us an overview of the energetic and temporal consequences of repeated stress. At last, technical and theoretical developments needed to carry this investigation further are presented.

7.7. Resistance to anti-cancer drugs by non-mutational mechanisms: insights from a cell-based multi-scale model of TRAIL-induced apoptosis

Participants: Virgile Andr eani, Gr egory Batt, Fran ois Bertaux.

The fact that tumors can acquire drug resistance by non-mutational mechanisms is increasingly gaining attention (Sharma et al., 2010; Pisco et al., 2013; Flusberg et al., 2013). Stochastic fluctuations in cellular states of different resistance and proliferative potential could play an important role in such resistance acquisition. Thus, to enable a quantitative, molecular-level understanding of those phenomena, modeling approaches that go beyond traditional, deterministic kinetic models of biological pathways are required.

An interesting and well-studied example of non-mutational resistance acquisition concerns the response of cancer cells to the agent TRAIL, a selective inducer of apoptotic cell death. In a previous work (Bertaux et al., 2014), we have developed a single-cell model of TRAIL-induced apoptosis that accounts for (1) protein-protein signaling reactions linking TRAIL exposure to commitment to apoptosis, (2) stochastic gene expression for the proteins involved in this signaling and (3) protein degradation. Under parsimonious and realistic assumptions for parameter values, fractional killing and transient resistance acquisition readily emerged from model simulations. Those two properties relating to TRAIL resistance are observed in-vitro for many different cancer cell lines.

Here, again in collaboration with Dirk Drasdo and Szymon Stoma, we investigate the long-term response of proliferating cancer cell populations repeatedly treated by TRAIL by integrating our single-cell model of TRAIL-induced apoptosis into a multi-cellular simulation framework. We predict that the long-term killing efficiency of repeated treatments is strongly reduced compared to the first treatment. A detailed analysis showed that resistance acquisition is caused mainly by the targeted degradation of activated pro-apoptotic proteins and an imbalance between the turnover of pro- and anti- apoptotic proteins. In addition, simulations of the treatment of multi-cellular spheroids suggested that limited TRAIL penetration is unlikely to be a driving cause of resistance, but that it can exacerbate the impact of cell-intrinsic resistance acquisition.

7.8. Controlling a genetic inverted pendulum

Participants: Gr egory Batt, Catherine Eisenhauer, Pascal Hersen, Jean-Baptiste Lugagne.

The ability to routinely control complex genetic circuits in vivo and in real-time promises quantitative understanding of cellular processes of unprecedented precision, quality, and richness. With combined efforts in microfluidic design, microscope automation, image segmentation and analysis, and control theory, we propose a platform for real-time, single-cell, externalized in silico control and monitoring of genetic networks in *E. coli*. Computational framework and hardware are optimized for parallelizing the experiments and we use the platform to test and control an entire library of synthetic genetic circuits. The circuits we are trying to control are based on the genetic toggle switch, a foundational circuit in synthetic biology, which consists of two genes that repress each other. This genetic system features two stable equilibrium points where one of the genes has taken over. Our objective is to dynamically balance the circuit in single cells around a third, unstable equilibrium point at which no gene dominates and their mutual repression strengths are balanced. This is similar to the landmark problem in control theory of stabilizing an inverted pendulum. Although our work indicates that this real-time control approach can drive convoluted genetic networks towards states that are inaccessible to traditional genetic perturbations such as knock-outs and promoter induction, the a priori quantitative knowledge of the system required for achieving this control is minimal. We show that even a simple Proportional-Integral controller can stabilize the unstable point of the toggle switch in single cells. Finally, we demonstrate that manipulation, or even inversion, of the stability map of the network is possible, though counter intuitive, via the simultaneous stabilization of an entire population of toggle switch cells around their unstable point with a common dynamic input.

7.9. Synthesizing Configurable Biochemical Implementation of Linear Systems from Their Transfer Function Specifications

Participants: Katherine Chiang, François Fages, Sylvain Soliman.

The ability to engineer synthetic systems in the biochemical context is constantly being improved and has a profound societal impact. Linear system design is one of the most pervasive methods applied in control tasks, and its biochemical realization has been proposed by Oishi and Klavins and advanced further in recent years. However, several technical issues remain unsolved. Specifically, the design process is not fully automated from specification at the transfer function level, systems once designed often lack dynamic adaptivity to environmental changes, matching rate constants of reactions is not always possible, and implementation may be approximative and greatly deviate from the specifications. In [6], building upon the work of Oishi and Klavins, we overcome these issues by introducing a design flow that transforms a transfer-function specification of a linear system into a set of chemical reactions, whose input-output response precisely conforms to the specification. This system is implementable using the DNA strand displacement technique. The underlying configurability is embedded into primitive components and template modules, and thus the entire system is adaptive. Simulation of DNA strand displacement implementation confirmed the feasibility and superiority of the proposed synthesis flow.

7.10. Reconfigurable Neuromorphic Computation in Biochemical Systems

Participants: Katherine Chiang, François Fages.

Implementing application-specific computation and control tasks within a biochemical system has been an important pursuit in synthetic biology. Most synthetic designs to date have focused on realizing systems of fixed functions using specifically engineered components, thus lacking flexibility to adapt to uncertain and dynamically-changing environments. To remedy this limitation, an analog and modularized approach to realize reconfigurable neuromorphic computation with biochemical reactions is presented in [11]. We propose a biochemical neural network consisting of neuronal modules and interconnects that are both reconfigurable through external or internal control over the concentrations of certain molecular species. Case studies on classification and machine learning applications using the DNA strand displacement technology demonstrate the effectiveness of our design in both reconfiguration and autonomous adaptation.

7.11. Search by Constraint Propagation

Participants: François Fages, Thierry Martinez, Sylvain Soliman.

Constraint programming is traditionally presented as the combination of two components: a constraint model and a search procedure. In [13] we show that tree search procedures can be fully internalized in the constraint model with a fixed enumeration strategy. This approach has several advantages: 1) it makes search strategies declarative, and modeled as constraint satisfaction problems; 2) it makes it possible to express search strategies in existing front-end modeling languages supporting reified constraints without any extension; 3) it opens up constraint propagation algorithms to search constraints and to the implementation of novel search procedures based on constraint propagation. We illustrate this approach with a Horn clause extension of the MiniZinc modeling language and the modeling in this language of a variety of search procedures, including dynamic symmetry breaking procedures and limited discrepancy search, as constraint satisfaction problems. We show that this generality does not come with a significant overhead, and can in fact exhibit exponential speedups over procedural implementations, thanks to the propagation of the search constraints.

7.12. Execution models for Constraint Programming and Semantics Equivalence

Participants: François Fages, Thierry Martinez, Sylvain Soliman.

Logic programming and constraint programming are two declarative programming paradigms which rely on the identification of programs to theories, and programming to modeling. Execution models result from the operational interpretation of logical provability in logic programming, and of constraint propagation in constraint programming. However, the control of execution is crucial for the practicability of these schemes and extra-logical traits are thus added in those programming systems, with the classical slogans "logic program = logical theory + control", "constraint program = constraint model + search".

In his thesis [4], Thierry Martinez investigates execution models in which control and search can be shifted into the logic or the constraint model, while preserving the semantics. The three parts of the thesis correspond to the three semantics equivalence that are showed: the first between two committed-choice forward-chaining logic languages, the second between constraint logic programs and constraint models, and the third between guard semantics in angelic settings. Each of these equivalence is constructive in the sense that there exists an encoding that enables the compilation from one of the paradigm to the other.

7.13. On Translating MiniZinc Constraint Model into Fitness Functions: Application to Continuous Placement Problems.

Participants: François Fages, Thierry Martinez, Bao Duy Tran.

MiniZinc is a solver-independent constraint modeling language which is increasingly used in the constraint programming community. It can be used to compare different solvers which are currently based on either constraint programming, Boolean satisfiability or mixed integer linear programming. In [12], we show how MiniZinc models can be compiled into fitness functions for evolutionary algorithms. More specifically, we describe the translation of FlatZinc models into fitness functions over the reals and their use in the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) solver. We illustrate this approach, and evaluate it, on the modeling and solving of complex shape continuous placement problems.

M3DISIM Team

7. New Results

7.1. Modeling

7.1.1. *Model-based analysis of continually measured signals of aortic pressure and flow*

Participants: Radomir Chabiniok, Dominique Chapelle [correspondant], Arthur Le Gall, Philippe Moireau, Fabrice Vallée.

We have started an application of reduced-order cardiac modeling in identifying relevant functional properties and state of heart from clinical records obtained during long-term (minutes-hours) monitoring of patients. Those are obtained either from anesthetized or intensive care patients by Fabrice Vallée, medical doctor in the department of anesthesia and intensive care at Lariboisière Hospital, Paris, who has joined the M3DISIM team in November 2015. The collaboration was initiated already in February 2015, and together with Fabrice we supervised the master's internship of Arthur Le Gall (medical doctor in his last year of specialization residency training). The internship took place at Lariboisière Hospital and in our lab at Inria Saclay (50:50%, period of April-September 2015). First published results are expected in 2016, when also a master's internship of a second student of Fabrice Vallée is scheduled. In addition, we intend to start a PhD on this topic in late 2016.

7.1.2. *Thermodynamical framework for modeling chemical-mechanical coupling in muscle contraction – Formulation and validation*

Participants: Matthieu Caruel, Dominique Chapelle [correspondant], Philippe Moireau.

Muscle contraction occurs at the nanoscale of a hierarchical multi-scale structure with the attachment of so-called cross-bridges within sarcomeres, namely, the creation of chemical bonds between myosin heads and specific sites on actin filaments. A cross-bridge in itself can be seen as a special chemical entity having internal mechanical variables – or degrees of freedom – pertaining to the actual geometric configuration, which implies that the free energy of the cross-bridge – whether in an attached or unattached state – must be made dependent on these internal variables (T.L. Hill, *Free Energy Transduction And Biochemical Cycle Kinetics*, Dover, 2004). This provides a thermodynamical basis for modeling the complex interplay of chemical and mechanical phenomena at the sarcomere level. Within this framework we propose a muscle model with two mechanical variables associated with a cross-bridge. For the action of individual cross-bridges occurring at the nanometer scale, the energy provided by the Langevin thermostat cannot be neglected, and we therefore propose to endow the internal mechanical variables with stochastic dynamics. Important motivations for this modeling choice include the ability to represent (i) the so-called power-stroke phenomenon and (ii) short-time responses of a muscle, e.g. to load steps. Our approach allows for systematic treatment of the model energetics, and in particular one goal of the proposed description is to investigate the potential benefit in mechanical efficiency with systems including – in addition to chemically-induced transformations – thermally-induced conformational changes such as the power-stroke.

7.1.3. *Biophysical modeling of seismocardiograms measurements*

Participants: Alexandre Laurin, Sébastien Imperiale [correspondant], Philippe Moireau, Dominique Chapelle.

We are developing models of various levels of complexity to represent seismocardiograms (SCG) that record mechanical thoracic vibrations induced by the beating heart. Our model combines a complete heartbeat model with a mechanical model of the thorax. The coupling is ensured by a unilateral contact modeling the non-penetration between the beating heart and the thoracic chest. In parallel, we are fine-tuning signal processing algorithms to identify the relevant characteristics of SCG and creating an iPhone application that is capable of acquiring the signal with its standard sensors. The application is also developed to integrate a simplified version of the cardio-thoracic model.

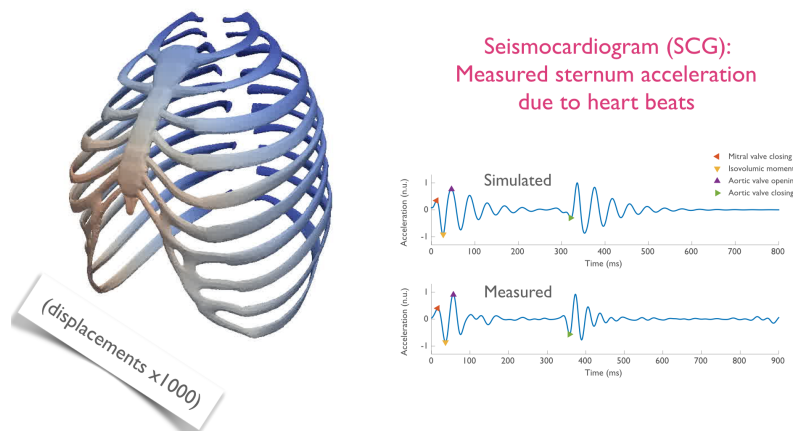


Figure 1. Model-based synthetic SCG compared to a measured SCG.

7.2. Numerical Analysis

7.2.1. Dirichlet-to-Neumann operator for diffraction problems in stratified anisotropic acoustic waveguides

Participants: Antoine Tonnoir [correspondant], Sonia Fliss [Poems team], Anne-Sophie Bonnet-Ben Dhia [Poems team].

In this work, we are interested in the construction of a Dirichlet-to-Neumann operator for the diffraction problem in stratified anisotropic acoustic waveguides. The key idea consists in using an adapted change of coordinates that allows to recover the completeness and the orthogonality of the modes on “deformed” cross-sections of the waveguide. Thus, we can properly define the diffraction problem and construct transparent boundary conditions to reformulate this problem in a bounded domain. Using classical arguments we easily prove the well-posedness. The method has also been implemented in a C++ code and has been validated.

7.2.2. Fourth-order energy-preserving locally implicit time discretization for linear wave equations

Participants: Sébastien Imperiale [correspondant], Juliette Chabassier [MAGIQUE-3D team].

In collaboration with Juliette Chabassier, we have constructed a family of fourth-order implicit-explicit time schemes for linear wave equations. Our application is the simulation of elastic waves propagation in a locally stiff medium. The domain of propagation is decomposed into several regions where different fourth-order time discretization are used, chosen among a family of implicit (for the stiff regions) or explicit fourth-order schemes. The coupling is based on a Lagrangian formulation on the boundaries between several non-conforming meshes of the regions. A global discrete energy is shown to be preserved and leads to global fourth-order consistency in time. Numerical results in 1D and 2D illustrate the good behavior of the schemes and their potential for the efficient simulation of realistic highly heterogeneous media for which using an explicit scheme everywhere can be extremely penalizing. Accuracy up to fourth-order reduces the numerical dispersion inherent to implicit methods used with a large time step, and makes this family of schemes attractive compared to second-order accurate methods.

7.2.3. Numerical methods for poromechanics: Applications to cardiac perfusion

Participants: Bruno Burtschell, Dominique Chapelle [correspondant], Philippe Moireau.

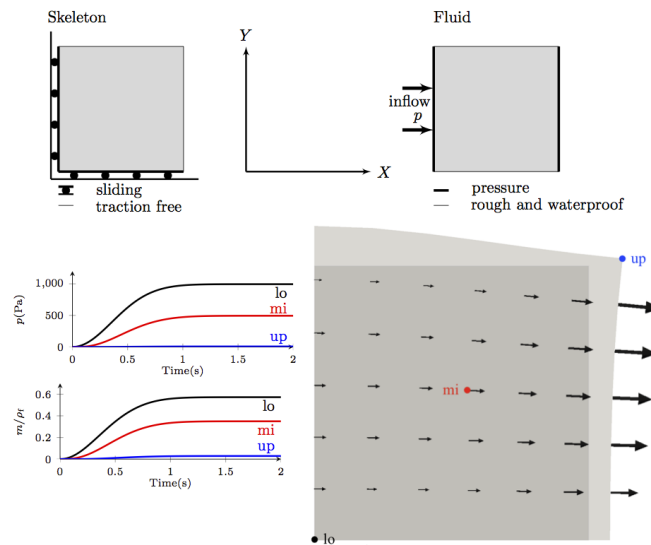


Figure 2. Swelling of a porous sample. On the left, fluid pressure and mass at three points in time during the swelling test. On the right, the fluid velocity field vector at the steady state in the deformed configuration.

We have previously formulated a rather general modeling framework of poromechanics – formulations that combine solid and fluid components to represent the behavior of a porous medium – to take into account large deformations and rapid fluid flows, see [6]. This allows to consider, in particular, the application of blood perfusion within the cardiac tissue, which features these specific complex phenomena, out of the scope of classical poromechanical models. One of our major objectives now, within the PhD of Bruno Burtschell, is to propose and analyse some associated relevant numerical schemes.

Some existing algorithms of fluid-structure interaction, with which our poromechanics formulations feature deep similarities, have been implemented – in FreeFEM++, both in axisymmetric configuration and in 3D – and compared. Their numerical and theoretical analysis – consistency, convergence – has been performed. Then, the adaptation of these algorithms to our poromechanics formulations enabled us to propose a time discretisation well-fitted to our framework, and to present its energy stability analysis. Spatial discretizations issues have also been specifically addressed, based on a complete analysis performed on a linearized problem, in order to guarantee pressure stability – via the selection of adequate inf-sup-compatible discretization spaces – including when the solid constituent is nearly or fully incompressible. Implementation and detailed numerical validations of these schemes have been performed. Integration into FELISCE (“HappyHeart” module) in 3D, and into a reduced model of cardiac cycle to take into account myocardium perfusion, are ongoing work.

7.3. Model-Data Interaction

7.3.1. Displacement Reconstructions in Ultrasound Elastography

Participant: Sébastien Imperiale.

In collaboration with Guillaume Bal (Columbia University, New York, USA), we have considered the reconstruction of internal elastic displacements from ultrasound measurements, which finds applications in the medical imaging modality called elastography. By appropriate interferometry and windowed Fourier transforms of the ultrasound measurements, we have proposed a reconstruction procedure of the vectorial

structure of spatially varying elastic displacements in biological tissues. This provides a modeling and generalization of scalar reconstruction procedures routinely used in elastography. The proposed algorithm has been justified using a single scattering approximation and local asymptotic analysis. Its validity has been assessed by numerical simulations.

7.3.2. *Recursive joint state and parameter estimation*

Participants: Atte Aalto, Philippe Moireau [correspondant].

We propose a method for estimating the parameters of a linear dynamical system from noisy measurements over a given, finite time, interval. For this purpose we develop a recursive modification of the joint state and parameter estimation method proposed in [7]. As the time interval is fixed, any errors in the initial state of the system may cause a significant error in the parameter estimate. Therefore, the parameter estimator is complemented by the so called back and forth nudging (BFN) method for estimating the system's initial state. The proposed strategy can also be regarded as a hybrid least squares optimization method for minimizing the quadratic discrepancy between the measured and simulated outputs over the set of all possible initial states and system parameters.

The optimality of the BFN method with colocated feedback has been considered as well. We have shown that in the case when the system's dynamics are governed by a skew-adjoint generator, the initial state estimate given by the BFN method converges to the minimizer of the quadratic output discrepancy – provided that the observer gains are chosen suitably. If the system's generator is essentially skew-adjoint and dissipative, a certain modification of the feedback operator is required in order to obtain such convergence.

7.3.3. *Convergence of discrete-time Kalman filter estimate to continuous-time estimate*

Participant: Atte Aalto [correspondant].

The Kalman(-Bucy) filter gives the optimal (minimum variance) solution to the state estimation problem for linear systems with Gaussian initial state, and white input and output noise processes. The implementation of the discrete-time Kalman filter is straightforward as it is readily formulated in an algorithmic manner. Thus, it may be tempting to use the discrete-time filter on the time-sampled continuous-time system. We study the convergence of the state estimate obtained from the discrete-time Kalman filter to the continuous-time estimate as the temporal discretization is refined. The convergence follows from the martingale convergence theorem, but surprisingly, no results exist on the rate of convergence. We derive convergence rate estimates for the discrete-time estimate under a number of different sets of assumptions starting from finite-dimensional systems and infinite-dimensional systems with bounded output operators and then proceeding to systems with unbounded output operators and systems with analytic semigroups. The proofs are based on applying the discrete-time Kalman filter on a dense, numerable subset of the time interval of interest, and bounding the change in the state estimate as the new data points are being added. These bounds, in turn, are based on smoothness estimates of the noise-free output.

7.3.4. *Observers for the wave equation in unbounded domains*

Participants: Sébastien Imperiale, Philippe Moireau [correspondant], Antoine Tonnoir, Sonia Fliss [Poems team], Karim Ramdani [Sphinx team].

We are interested in the reconstruction of initial data for the wave equation problem in unbounded domains using an observer strategy. A major advantage of this method for problems set in bounded domains is the exponential convergence of the algorithm of reconstruction. In our case, the specificity is the unboundedness of the domain which requires to bound it with artificial boundaries for numerical computations. To avoid spurious reflections due to these artificial boundaries, we consider transparent boundary conditions. The difficulty then is to adapt the classical observers technique to this case. Indeed, after enough time, the outgoing waves have left the computational domain and the related information is in some sense “lost”.

First results have been obtained for the 1D case: the theoretical proof of the (exponential) convergence of the algorithm has been done, and the method has been numerically validated. We are currently working on the extension to the 2D case, which raises new difficulties. In particular, the construction of the transparent boundary condition is not obvious and implies a non-local operator in both time and space. Due to this non-local operator, the theoretical analysis of the convergence of the method is then much more difficult.

7.3.5. A Luenberger observer for reaction-diffusion models with front position data

Participants: Dominique Chapelle, Annabelle Collin, Philippe Moireau [correspondant].

We propose a Luenberger observer for reaction-diffusion models with propagating front features, and for data associated with the location of the front over time. Such models are considered in various application fields, such as electrophysiology, wild-land fire propagation and tumor growth modeling. Drawing our inspiration from image processing methods by considering a data similarity measure of Mumford-Shah type, we start by proposing an observer for the eikonal-curvature equation that can be derived from the reaction-diffusion model by an asymptotic expansion. We then carry over this observer to the underlying reaction-diffusion equation by an “inverse asymptotic analysis”, and we show that the associated correction in the dynamics has a stabilizing effect for the linearized estimation error. We also discuss the extension to joint state-parameter estimation by using the earlier-proposed ROUKF strategy. We published a first work [17] where the observer feedback is derived from the shape-derivative of the data similarity measure. Then, in [21], in order to improve the observer formulation, we followed a strategy of increasing importance in shape optimization or “level-set”-based image segmentation by complementing the required shape derivatives, used to modify the shape contours, by a topological derivative that represents the sensitivity of the similarity measure when removing a small part of the domain. Both results are illustrated with test problems pertaining to electrophysiology modeling, including with a realistic model of cardiac atria. Our numerical trials show that state estimation is directly very effective with the proposed Luenberger observer.

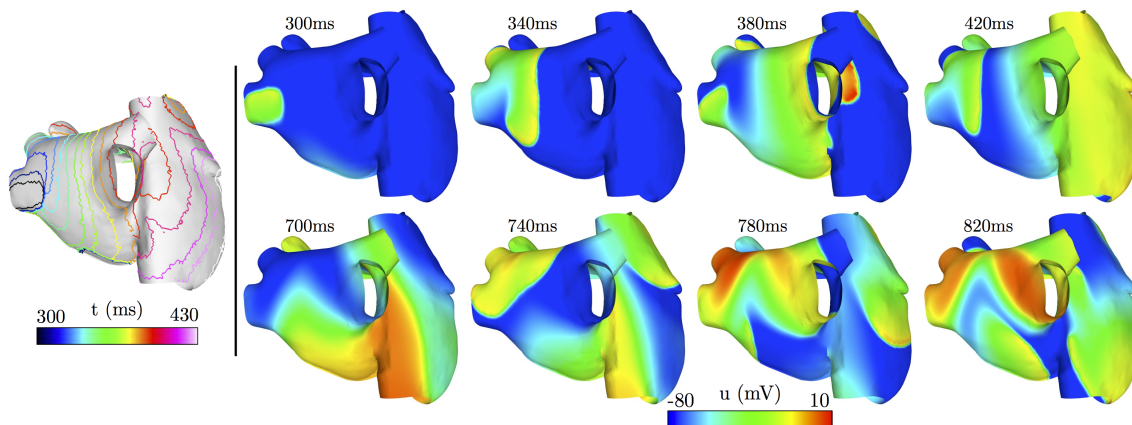


Figure 3. Atrial fibrillation: synthetic front data with noise (left, note that only the first times of passage after the onset of fibrillation are displayed, for the sake of clarity), and corresponding observer solutions (right)

7.3.6. Identification of weakly coupled multiphysics problems. Application to the inverse problem of electrocardiography

Participants: Cesare Corrado [Reo team], Jean-Frédéric Gerbeau [Reo team], Philippe Moireau [correspondant].

This work addresses the inverse problem of electrocardiography from a new perspective, by combining electrical and mechanical measurements. Our strategy relies on the definition of a model of the electromechanical contraction which is registered on ECG data, but also on measured mechanical displacements of the heart tissue typically extracted from medical images. In this respect, we establish in this work the convergence of a sequential estimator which combines for such coupled problems various state-of-the-art sequential data assimilation methods in a unified consistent and efficient framework. Indeed, we aggregate a Luenberger observer for the mechanical state and a Reduced-Order Unscented Kalman Filter applied on the parameters to be identified and a POD projection of the electrical state. Then, using synthetic data we show the benefits of our approach for the estimation of the electrical state of the ventricles along the heart beat, compared with more classical strategies that only consider an electrophysiological model with ECG measurements. Our numerical results actually show that the mechanical measurements improve the identifiability of the electrical problem, allowing to reconstruct the electrical state of the coupled system more precisely. Therefore, this work is intended to be a first proof of concept, with theoretical justifications and numerical investigations, of the advantage of using available multi-modal observations for the estimation and identification of an electromechanical model of the heart.

7.3.7. Data assimilation of cine-MR images by a biophysical model

Participants: Radomir Chabiniok, Dominique Chapelle [correspondant], Alexandra Groth, Philippe Moireau, Juergen Weese.

Within the European project VP2HF, we participated in extending the image segmentation tool developed by Philips Hamburg (Alexandra Groth, Jürgen Weese) to process clinically routine cine-MR images for creating anatomical models of heart. Secondly, together with A. Groth and J. Weese we defined a discrepancy operator – between a biomechanical heart model and cine-MR images – that does not require segmenting MR images prior to data assimilation. Initial results of the state estimation using this discrepancy operator were presented at the 2nd VP2HF evaluation meeting (December 2015), and extending these results into a journal paper is a joint objective of the M3DISIM team and of Philips Hamburg.

MAGIQUE-3D Project-Team

6. New Results

6.1. Seismic Imaging and Inverse Problems

6.1.1. *hp-adaptive simulation and inversion of magnetotelluric measurements*

Participants: Hélène Barucq, Julien Alvarez Aramberri, David Pardo.

The magnetotelluric (MT) method is a passive exploration technique that aims at estimating the resistivity distribution of the Earth's subsurface, and therefore at providing an image of it. This process is divided into two different steps. The first one consists in recording the data. In a second step, recorded measurements are analyzed by employing numerical methods. In this work, we provide a rigorous mathematical setting in the context of the Finite Element Method (FEM) that helps to understand the MT problem and its inversion process. In order to recover a map of the subsurface based on 2D MT measurements, we employ for the first time in MTs a multigoal oriented self adaptive hp-Finite Element Method (FEM). We accurately solve both the full formulation as well as a secondary field formulation where the primary field is given by the solution of a 1D layered medium. To truncate the computational domain, we design a Perfectly Matched Layer (PML) that automatically adapts to high-contrast material properties that appear within the subsurface and on the air-ground interface. For the inversion process, we develop a first step of a Dimensionally Adaptive Method (DAM) by considering the dimension of the problem as a variable in the inversion. Additionally, this dissertation supplies a rigorous numerical analysis for the forward and inverse problems. Regarding the forward modelization, we perform a frequency sensitivity analysis, we study the effect of the source, the convergence of the hp-adaptivity, or the effect of the PML in the computation of the electromagnetic fields and impedance. As far as the inversion is concerned, we study the impact of the selected variable for the inversion process, the different information that each mode provides, and the gains of the DAM approach.

6.1.2. *Ultrasonic imaging of complex media*

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé, Julien Diaz, Sébastien Tordeux, Ha Howard Faucher.

In 2015 we have begun a collaborating project with I2M (Physics Acoustics Department of Bordeaux 1 University). We aim at modeling and simulating efficiently the propagation of acoustic waves and later elastodynamic waves in highly heterogeneous media, the final goal is to use topological gradient imaging techniques. Classical techniques as finite elements can be too costly, we propose to design more efficient numerical techniques that exploit the fact that the wavelength is big with respect to the heterogeneities. For instance, we will use numerical upscaling, multiscale homogenization or asymptotic methods. A funding has been obtained for a PhD and a post doctoral position, that have both started in 2015. Our first step is to design a laboratory experiment and a simulation code in order to challenge the limits of the newly derived models and quantify their validity.

6.1.3. *Impedance transmission conditions for the electric potential across a highly conductive casing*

Participants: Hélène Barucq, Aralar Erdozain, David Pardo, Victor Péron.

In this study we present Impedance Transmission Conditions (ITCs) for the electric potential in the framework of borehole through-casing resistivity measurements. Such ITCs substitute the part of the domain corresponding to a highly conductive casing. The naturally small thickness of the casing makes it ideal for exhibiting ITCs. We numerically observe the delivered order of accuracy.

6.1.4. An efficient truncated SVD of large matrices based on the low-rank approximation for inverse geophysical problems

Participant: Sébastien Tordeux.

We have proposed a new algorithm to compute a truncated singular value decomposition of the Born matrix based on a low-rank arithmetic. Theoretical background to the low-rank SVD method has been investigated: the Born matrix of an acoustic problem can be approximated by a low-rank approximation derived thanks to a kernel independent multipole expansion. The new algorithm to compute T-SVD approximation consists of four steps, and they are described in detail. The largest singular values and their left and right singular vectors can be approximated numerically without performing any operation with the full matrix. The low-rank approximation is computed due to a dynamic panel strategy of cross approximation technique.

6.1.5. Handling clusters with a task-based runtime system: application to Geophysics

Participants: Emmanuel Agullo, Hélène Barucq, Lionel Boillot, George Bosilca, Julien Diaz.

The extreme complexity of hardware platforms makes them harder and harder to program. To fully exploit such machines, the High Performance Computing community often uses a MPI + X (X being pthreads, OpenMP, Cuda ...) programming models. We propose to use an alternative solution consisting of programming at a higher level of abstractions by describing a scientific, high performance computing application as a sequence of tasks whose execution is delegated to a runtime system. We compared MPI-based version and task-based version on Geophysics simulations, especially on the DIVA code of Total. Our previous studies demonstrated the task-based paradigm superiority on shared memory architectures (CPU or MIC), we are now working on distributed and heterogeneous architectures (CPUs+MICs) and, according to our preliminary results, the performances are still better than the MPI-version.

This work has been presented to the conferences PRACEdays [60], Rice Oil&Gas [43] and PASC [37].

6.2. Mathematical modeling of multi-physics involving wave equations

6.2.1. Elasto-acoustic coupling

Participants: Hélène Barucq, Lionel Boillot, Henri Calandra, Julien Diaz, Simon Ettouati, Conrad Hillairet, Elvira Shishenina.

In the framework of her Master thesis, Elvira Shishenina developed a Discontinuous Galerkin Method for the elastoacoustic coupling in time domain. The proposed solution methodology in general and can be applied to any kind of fluxes. We have implemented and validated in Elasticus a centered flux version and an upwind flux version in two dimensions. The time discretization is achieved thanks to Runge Kutta schemes of second and fourth orders.

In frequency domain, Conrad Hillairet developed a 3D elasto-coupling IPDG scheme, in the framework of his Master thesis. It has been implemented and validated in Hou10ni. Moreover, the code is able to handle p -adaptivity and we have proposed a strategy in order to determine the order of the cell as a function of the size of the cell and of the physical parameters. The results of this work have been presented to the Siam Conference on Geosciences in Stanford [39] and to the XXIV Congress on Differential Equations and Applications in Cadiz [32].

Finally, we have considered elastoacoustic coupling with curved interfaces and we have proposed a solution methodology based on Finite Element techniques, which allows for a flexible coupling between the fluid and the solid domain by using non-conforming meshes and curved elements. Since characteristic waves travel at different speeds through different media, specific levels of granularity for the mesh discretization are required on each domain, making impractical a possible conforming coupling in between. Advantageously, physical domains may be independently discretized in our framework due to the non-conforming feature. Consequently, an important increase in computational efficiency may be achieved compared to other implementations based on non-conforming techniques, namely by reducing the total number of degrees of freedom. Differently from other non-conforming approaches proposed so far, our technique is relatively simpler and requires only a

geometrical adjustment at the coupling interface at a preprocessing stage, so that no extra computations are necessary during the time evolution of the simulation. On the other hand, as an advantage of using curvilinear elements, the geometry of the coupling interface between the two media of interest is faithfully represented up to the order of the scheme used. In other words, higher order schemes are in consonance with higher order approximations of the geometry. Concerning the time discretization, we analyzed both explicit and implicit schemes. These schemes are energy conserving and, for the explicit case, the stability is guaranteed by a CFL condition.

This work, which has been achieved in collaboration with Angel Rodriguez Rozas, former post-doc of the team, was published in *Journal of Computational Physics* [27].

6.2.2. *Atmospheric boundary conditions for helioseismology*

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé, Victor Péron.

The sun does not have a clear boundary like a solid ball, but it has an atmosphere which can be modeled as an exponential decay of the density. We have studied the replacement of this atmosphere by an equivalent boundary condition in order to avoid meshing the atmosphere. When we assume that the exponential decay is large enough, asymptotic modeling can be performed with respect to this large parameter. Equivalent boundary conditions have been obtained for order 1, 2 and 3, and they substantially improve Dirichlet condition (order 0) for low frequencies. However for high frequencies, these conditions are no longer relevant. We have developed a first-order absorbing boundary condition adapted to an exponential decay of the density, this last condition provides good results for the tested range of frequency. These conditions have been used by the team of Laurent Gizon (Max Planck institute) to retrieve experimental dispersion curves, so called “power spectrum”.

6.2.3. *Absorbing Boundary Conditions for 3D elastic TTI modeling*

Participants: Hélène Barucq, Lionel Boillot, Julien Diaz.

We propose stable low-order Absorbing Boundary Conditions (ABC) for elastic TTI modeling. Their derivation is justified in elliptic TTI media but it turns out that they are directly usable to non-elliptic TTI configurations. Numerical experiments are performed by using a new elastic tensor source formula which generates P-waves only in an elliptic TTI medium. Numerical results have been performed in 3D to illustrate the performance of the ABCs.

This work has been presented to the conferences PANACM [38] and SEG [33].

6.2.4. *The airfoil equation on near disjoint intervals : Approximate models and polynomial solutions*

Participants: Leandro Farina, Marcos Ferreira, Victor Péron.

In [26], the airfoil equation is considered over two disjoint intervals. Assuming the distance between the intervals is small an approximate solution is found and relationships between this approximation and the solution of the classical airfoil equation are obtained. Numerical results show the convergence of the approximation to the solution of the original problem. Polynomial solutions for an approximate model are obtained and a spectral method for the generalized airfoil equation on near disjoint intervals is proposed.

6.2.5. *Finite element subproblem method*

Participants: Patrick Dular, Christophe Geuzaine, Laurent Krähenbühl, Victor Péron.

In [25], progressive refinements of inductors are done with a subproblem method, from their wire or filament representations with Biot-Savart models up to their volume finite-element models, from statics to dynamics. The reaction fields of additional magnetic and/or conducting regions are also considered. Accuracy improvements are efficiently obtained for local fields and global quantities, i.e., inductances, resistances, Joule losses, and forces.

6.2.6. Asymptotic study for Stokes-Brinkman model with Jump embedded transmission conditions

Participants: Philippe Angot, Gilles Carbou, Victor Péron.

In [18], one considers the coupling of a Brinkman model and Stokes equations with jump embedded transmission conditions. Assuming that the viscosity in the porous region is very small, we derive a Wentzel-Kramers-Brillouin (WKB) expansion in power series of the square root of this small parameter for the velocity and the pressure which are solution of the transmission problem. This WKB expansion is justified rigorously by proving uniform errors estimates.

6.2.7. On the solution of the Laplace equation in 3-D domains with cracks and elliptical edges

Participants: Victor Péron, Samuel Shannon, Zohar Yosibash.

An explicit asymptotic solution to the elasticity system in a three-dimensional domain in the vicinity of an elliptical crack front, or for an elliptical sharp V-notch is still unavailable. Towards its derivation we first consider in [30] the explicit asymptotic solutions of the Laplace equation in the vicinity of an elliptical singular edge in a three-dimensional domain. Both homogeneous Dirichlet and Neumann boundary conditions on the surfaces intersecting at the elliptical edge are considered. The dual singular solution is also provided to be used in a future study to extract the edges flux intensity functions by the quasi-dual function method. We show that just as for the circular edge case, the solution in the vicinity of an elliptical edge is composed of three series, with eigenfunctions being functions of two coordinates.

In [29] the singular solution of the Laplace equation with a straight-crack is represented by a series of eigenpairs, shadows and their associated edge flux intensity functions (EFIFs). We address the computation of the EFIFs associated with the integer eigenvalues by the quasi dual function method (QDFM). The QDFM is based on the dual eigenpairs and shadows, and we show that the dual shadows associated with the integer eigenvalues contain logarithmic terms. These are then used with the QDFM to extract EFIFs from p-version finite element solutions. Numerical examples are provided.

6.3. Supercomputing for Helmholtz problems

6.3.1. High order methods for Helmholtz problems in highly heterogeneous media

Participants: Théophile Chaumont-Frelet, Henri Calandra, H el ene Barucq, Christian Gout.

The numerical solution of Helmholtz problems set in highly heterogeneous media is a tricky task. Classical high order discretizations fail to handle such propagation media, because they are not able to capture any of the scales of the velocity parameter. Indeed, they are build upon coarse meshes and therefore, if the velocity parameter is taken to be constant in each cell (through averaging, or local homogenization strategy), scale information is (at least partially) lost. We propose to overcome this difficulty by introducing a multiscale medium approximation strategy. The velocity parameter is not assumed to be constant on each cell, but on a submesh of each cell. If the submeshes are designed properly, the medium approximation method is equivalent to a quadrature formula, adapted to the medium. In particular, we show that this methodology has roughly the same computational cost as the classical finite element method. This new solution methodology has been presented in a paper under revision. We have performed a mathematical analysis of the multiscale medium approximation techniques to higher order discretization. First, we show that the heterogeneous Helmholtz problem is well-posed and derive stability estimates with respect to the right hand side, and with respect to variations of the velocity parameter, justifying the use of medium approximation. Those results are obtained assuming the velocity parameter is monotonous and that the propagation medium is closed by first order absorbing boundary conditions. However, these hypothesis are not mandatory to discretize the problem. Second, we turn to the analysis of finite element schemes with subcell variations of the velocity. In particular, we show that even if the solution can be rough inside each cell because of velocity jumps, we are able to extend the asymptotic error estimates obtained in [93] to heterogeneous media with non-matching mesh in case of elements of order $1 \leq p \leq 3$. Third, we investigate numerically the stability of the scheme when the

frequency is increasing to figure out optimal meshing conditions. We show that in simple media, the optimal homogeneous pre-asymptotic error estimates are still valid. However, in more complex cases, it looks like this condition is not sufficient anymore. Apart from showing that the homogeneous results are not always applicable to the heterogeneous Helmholtz equation, we are not able to give a clear answer to the question. Finally, we are able to conclude that high order methods are actually interesting: in our examples, $p = 4$ discretizations always yield a smaller linear system than lower order discretizations for the same precision.

6.3.2. Hybridizable Discontinuous Galerkin method for the elastic Helmholtz equations

Participants: Marie Bonnasse-Gahot, Henri Calandra, Julien Diaz, Stéphane Lanteri.

In the framework of the PhD thesis of Marie Bonnasse-Gahot, we have proposed an hybridizable discontinuous Galerkin method for solving the anisotropic elastodynamics wave equations in harmonic domain, in two and three dimensions. The method was implemented in Hou10ni and in the platform of Total. We have analyzed the performance of the proposed method in 2D on simple test case and compared it to classical DG methods. We have shown that the HDG method provides a more accurate solution for less computational cost provided that the order is high enough. We have illustrated the usefulness of the p -adaptivity in 2D, which allows to reach the accuracy of a global method of degree p for the costs of a global method of degree $p - 1$ or $p - 2$. This feature is already implemented in the 3D code. We now have to determine an accuracy criteria for assigning an order to a given cell, similar to the criteria we proposed in 2D.

For the numerical analysis of the scheme, we have shown that the HDG method could be rewritten as an upwind fluxes DG method and one of our perspectives is to use this equivalence in order to perform a dispersion analysis following the work of Ainsworth, Monk and Muniz [64].

We have shown that HDG could be used for 2D simulation on geophysical benchmark, and we will now implement the method in a Reverse Time Migration software, the ultimate goal being to couple HDG method with a full waveform inversion solver. In order to tackle more realistic test cases in 3D, it will be mandatory to improve the linear solver and we are now considering the use of an hybrid solver such as Maphys developed by the Inria team-project HIEPACS.

The results of this work have been presented at the ‘‘SIAM Conference on Geosciences’’ [48] and at the ‘‘Oil and Gas HPC Workshop’’ [49].

6.4. Hybrid time discretizations of high-order

6.4.1. High-order symmetric multistep schemes for wave equation

Participants: Juliette Chabassier, Marc Durufl e, Guillaume Marty.

We have studied high-order symmetric multistep schemes for the second-order formulation $y'' = f(t, y)$ during the internship of Guillaume Marty. The stability condition (CFL) can be optimized for explicit schemes since they have free parameters. However, this optimization procedure is not easy since the optimum is reached for forbidden values (values for which the high-order accuracy is no longer obtained). We have proposed acceptable values of free parameters for schemes of order 4, 6 and 8. These schemes have been tested for the wave equation, they suffer from a lack of robustness with respect to rounding numerical errors. The stability of implicit schemes has also been explored. For fourth-order schemes, a family of energy-conserving schemes has been obtained. However, we have not found unconditionally stable high-order schemes, which is well-known for the first-order formulation as Dahlquist’s barrier. It seems that for the second-order formulation, this barrier holds and only second-order accurate schemes are unconditionally stable. Implicit high-order schemes have a maximum CFL of $\sqrt{6}$, the same CFL as the standard θ -scheme with $\theta = \frac{1}{12}$. As a result, the implicit version of these schemes does not have a practical interest.

6.4.2. High order conservative explicit and implicit schemes for wave equations.

Participants: Juliette Chabassier, S ebastien Imperiale.

In 2015 we have studied the space/time convergence of a family of high order conservative explicit and implicit schemes for wave equations. An original proof of convergence has been proposed and provides an understanding of the lack of convergence of some schemes when the time step approaches its greatest admissible value for stability (CFL condition). An article is being written and will be submitted soon.

6.4.3. Multi-level explicit local time-stepping methods for second-order wave equations

Participants: Julien Diaz, Marcus Grote.

Local mesh refinement severely impedes the efficiency of explicit time-stepping methods for numerical wave propagation. Local time-stepping (LTS) methods overcome the bottleneck due to a few small elements by allowing smaller time-steps precisely where those elements are located. Yet when the region of local mesh refinement itself contains a sub-region of even smaller elements, any local time-step again will be overly restricted. To remedy the repeated bottleneck caused by hierarchical mesh refinement, multi-level local time-stepping methods are proposed, which permit the use of the appropriate time-step at every level of mesh refinement. Based on the LTS methods from Diaz and Grote [82], these multi-level LTS methods are explicit, yield arbitrarily high accuracy and conserve the energy.

The method was published in *Computer Methods in Applied Mechanics and Engineering* [24].

MAMBA Project-Team

7. New Results

7.1. Cancer

Participants: Luís Lopes Neves de Almeida, Rebecca Chisholm, Jean Clairambault, François Delhommeau [Haematology department, St Antoine Hospital, Paris], Dirk Drasdo, Ján Eliaš, Alexandre Escargueil [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Ghassen Haddad [ENIT, Tunis], Shalla Hanson [Department of mathematics, Duke University, Durham, NC], Pierre Hirsch [Haematology department, St Antoine Hospital, Paris], Groups Invade, Lungsysii, Tim Johann, Group Klingmueller [German Cancer Center, Heidelberg], Michal Kowalczyk [Univ. Santiago de Chile], Annette Larsen [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Tommaso Lorenzi, Alexander Lorz, Benoît Perthame, Andrada Quillas Maran, Fernando Quirós [Univ. Autónoma de Madrid], Michèle Sabbah [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Min Tang [Jiaotong University, Shanghai], Emmanuel Trélat [LJLL, UPMC], Paul Van Liedekerke, Nicolas Vauchelet, Irène Vignon-Clementel [REO], Yi Yin.

7.1.1. Drug resistance

We have continued to develop our phenotypically based models of drug-induced drug resistance in cancer cell populations, representing their Darwinian or Lamarckian evolution under drug pressure by integro-differential equations. In one of them [23], a 1D space variable has been added to the phenotypic structure variable to account for drug diffusion in tumour spheroids. In another one, focusing on both Darwinian selection and Lamarckian-like (non-genetic) instruction, published in *Cancer Research* [41], where deterministic and agent-based modelling are processed in parallel, we have added advection and diffusion terms to the initial integro-differential model and considered a physiologically based 2-dimensional phenotypic structure variable. This model, designed to take account of previously published biological observations on (reversible) drug tolerance persistence in a cultured population of non-small cell lung cancer (NSCLC) cells [90], reproduces the observations and we propose to assess the model by testing biologically based hypotheses. This work, also presented in various conferences ([34], [35], [31]) is conducted in close collaboration with the INSERM-UPMC team “Cancer biology and therapeutics” (A. Larsen, A. Escargueil, M. Sabbah) at St Antoine Hospital. It has also led our postdoctoral fellows Rebecca Chisholm and Tommaso Lorenzi to prolong their work on the *Cancer Research* paper by publishing two more articles [21], [48], one of which is a joint work with Alexander Lorz. This work is currently continued from the point of view of optimal control in Camille Pouchol’s PhD thesis.

7.1.2. Evolution and cancer, therapy optimisation

Guided by our goal to understand and overcome drug resistance in cancer cell populations[41], we are considering cancer as an evolutionary phenomenon at two time scales: a large time scale (billions of years) of evolution of the genomes, from unicellular organisms to organised multicellularity (viewing cancer as more an archeoplasm than a neoplasm, an evolution backwards, following Davies and Lineweaver, *Phys Biol* 2011, and others [78], [66], [92], [79]) with shortcomings due to malfunctions in the processes of control of cell differentiation, and a short time scale (duration of a human life) of evolution in the “epigenetic landscape” of a given genome (as advocated by Sui Huang and Angela Pisco, e.g. recently in *Nature*, *Br J Cancer* and elsewhere [76], [77], [85], [86], [94]). It leads us to propose theoretical frameworks for innovative cancer therapeutics from this evolutionary biology viewpoint, taking into account the major clinical issue of drug resistance in cancer cell populations, as presented in [31] and exposed to a medical audience at the symposium “Réseau Cancer des Points Cardinaux” (http://www.frog-oncogeriatric.com/fichiers/evnmt_41.pdf).

7.1.3. Interactions between tumour cell populations and their cellular micro-environments

A phenotype-structured model of the interactions between a breast cancer cell population (MCF7 cultured cells, collaboration with M. Sabbah, St Antoine Hospital) and its adipocyte stroma support cell population has been developed (T. Lorenzi, C. Pouchol, J. Clairambault) in the framework of Camille Pouchol's Inria internship ([56]). It has led to hiring C. Pouchol as a PhD student at UPMC (on a university grant "Interfaces pour le Vivant") on the same subject with perspectives in optimal therapeutic control, under the supervision of J. Clairambault, M. Sabbah and E. Trélat, see below "Supervision".

7.1.4. Combining chemo- and immunotherapies

Both from the point of view of interactions with the tumour micro-environment and of innovative anticancer therapies, it is necessary to take into account the immune response in cancer. This recently developed activity, (illustrated by presentations in session 70 in ICNAAM 2016 [32]) has led to the involvement in 2015 of Shalla Hanson as a PhD student in co-tutela between Duke University, NC and UPMC, see below "Supervision".

7.1.5. Hele-Shaw model of tumour growth

The mathematical analysis of macroscopic models of tumor growth with one type of cancer cells has been continued. On the one hand, in [47], the concept of viscosity solutions has been implemented for the case with active motion. On the other hand, the regularity of the free boundary is proved in [51] using methods developed for the standard Hele-Shaw equation and a new formulation.

7.1.6. The p53 protein spatio-temporal dynamics

Our previously developed spatio-temporal models for an intracellular dynamical response of the p53 protein to DNA damage, have been exploited further, and several testable biological hypotheses have been proposed in [33]. Among them, we suggest ideas that link spatio-temporal location of the p53 protein with a specific cell fate of a single cell in [33], [2] and, based on our new oscillator relying on both positive and negative regulation of p53 by Mdm2 (in tight cooperation with MdmX), we provide molecular insights into an excitability of the p53 network, i.e., we propose a molecular explanation for a full pulsatile response of p53 independently of input (ATM) signalling, challenging thus different fates of ATM downstream targets in the regulation of p53 in response to different stimuli, such as γ - and UV-radiation.

Our mathematical models, all included in J. Eliaš's PhD thesis [2] (defended on the 1st September 2015), contribute to understanding the variability of p53 in response to single and double strand breaks and reveal some new aspects of the core p53-Mdm2 protein feedback.

7.1.7. Lung and breast cancer

We developed an image analysis software and designed image analysis pipelines which we used to quantify the invasion pattern of non-small cell lung cancer (NSCLC) cells in multicellular spheroid *in vitro* experiments [24]. Based on the analyses, we demonstrated that the concomitant over-expression of FIR (far upstream element binding protein interacting repressor) and its splice variants drives NSCLC migration and dissemination.

We developed an agent-based, centre-based model of cell migration in cancer invasion based upon experimental observations of cell shape and cell behaviour in multicellular spheroid experiments of breast and lung cancer cells. In these experiments, cells deform from a sphere into an oblong shape upon migration, and adopt a spherical shape again whenever they turn back to such spheroids. This was implemented. Moreover, we developed a 3D model for the extracellular matrix (ECM) in which the matrix is modelled by an irregular network of springs with nodes represented as elastic objects. Migrating cells anchor in the network to move, leading to network deformation. We implemented a number of different biological mechanisms of cell migration and cell-ECM interaction. We find that a relatively simple model is sufficient to explain all phenomena of a single invading cell (Palm et. al., in preparation).

The combination of image analysis and of the abovementioned refined invasion model should allow a quantitative model of multicellular invasion following the same line of research as for SK-MES-1 cells, where we inferred a multicellular spheroid growth model from image data within a pipeline of experiment, imaging, image analysis and modelling [17]. In that paper, we used spatial-temporal image data of cell nucleus distribution, cell proliferation, death, and ECM distribution for two growth conditions (oxygen and glucose) to calibrate a model which was then able to quantitatively correctly predict the growth kinetics of the tumor spheroids for two other growth conditions, one strongly glucose limited, another strongly oxygen-limited.

Finally, we developed an image analysis pipeline to estimate the number of cancer cells in a patient with non-small cell lung cancer (NSCLC) from non-invasive image modalities. The estimate bases upon cell counts from histological serial sections of the tumor which have been related to the D-value inferred from Diffusion Weighted (DW) MRI (Yi et. al., paper in preparation).

7.2. Aggregation Kinetics

Participants: Aurora Armiento, Tom Banks [CRSC, NCSU, Raleigh, USA], Thibault Bourgeron, José Antonio Carrillo [Imperial College, London, United Kingdom], Marie Doumic, Miguel Escobedo [Universidad del País Vasco, Bilbao, Spain], Sarah Eugène, Marc Hoffmann [Ceremade, Université Paris-Dauphine], François James [MAPMO, Université d'Orléans], Nathalie Krell [Université de Rennes 1], Carola Kruse, Frédéric Lagoutière [Département de mathématiques d'Orsay], Philippe Moireau [Inria Paris Saclay, M3DISIM project-team], Benoît Perthame, Stéphanie Prigent, Human Rezaei [VIM, INRA Jouy-en-Josas], Lydia Robert [Laboratoire Jean Perrin, UPMC], Philippe Robert [Inria Paris, RAP project-team], Maria Teresa Teixeira [IBCP, Paris], Nicolas Vauchelet, Min Tang [Jiaotong University, Shanghai], Zhou Xu [IBCP, Paris], Wei-Feng Xue [University of Kent, United Kingdom].

7.2.1. Heterogeneity as an intrinsic feature in biological dynamics

Combining deterministic and probabilistic approaches, we investigated in two different applications - namely senescence and protein aggregation - the impact of heterogeneity on dynamical features of the considered populations.

Yeast Senescence and Telomere replication In eukaryotes, the absence of telomerase results in telomere shortening, eventually leading to replicative senescence, an arrested state that prevents further cell divisions. While replicative senescence is mainly controlled by telomere length, the heterogeneity of its onset is not well understood. Insights on this key question may have consequences both for cancer and aging issues.

In collaboration with T. Teixeira and Z. Xue from IBCP, we proposed a mathematical model based on the molecular mechanisms of telomere replication and shortening to decipher the causes of this heterogeneity [7]. Using simulations fitted on experimental data obtained from individual lineages of senescent *Saccharomyces cerevisiae* cells, we decompose the sources of senescence heterogeneity into interclonal and intracolonial components, and show that the latter is based on the asymmetry of the telomere replication mechanism. We also evidence telomere rank-switching events with distinct frequencies in short-lived versus long-lived lineages, revealing that telomere shortening dynamics display important variations. Thus, the intrinsic heterogeneity of replicative senescence and its consequences find their roots in the asymmetric structure of telomeres.

These promising first results lead us to an ongoing collaboration, and hopefully will allow still more insight on complex mechanisms not yet modelled mathematically.

Variability in nucleated polymerisation

The kinetics of amyloid assembly show an exponential growth phase preceded by a lag phase, variable in duration as seen in bulk experiments and experiments that mimic the small volumes of cells. To investigate the origins and the properties of the observed variability in the lag phase of amyloid assembly currently not accounted for by deterministic nucleation dependent mechanisms, we formulated a new stochastic minimal model that is capable of describing the characteristics of amyloid growth curves despite its simplicity [44]. We then solved the stochastic differential equations of our model and gave a mathematical proof of a central limit theorem for the sample growth trajectories of the nucleated aggregation process. These results

give an asymptotic description for our simple model, from which closed-form analytical results capable of describing and predicting the variability of nucleated amyloid assembly were derived. We also demonstrated the application of our results to inform experiments in a convenient and clear way. Our model offers a new perspective and paves the way for a new and efficient approach on extracting vital information regarding the key initial events of amyloid formation.

7.2.2. Inverse Problems and Data Assimilation Applied to Protein Aggregation and other settings

As mathematical models become more complex with multiple states and many parameters to be estimated using experimental data, there is a need for critical analysis in model validation related to the reliability of parameter estimates obtained in model fitting. This leads to a fundamental question: how much information with respect to model validation can be expected in a given data set or collection of data sets?

In the biological context of amyloid formation, the question is to quantify to which extent a given model may be appropriately fitted and selected for, given relatively sparse data. Estimating reaction rates and size distributions of protein polymers is an important step towards understanding the mechanisms of protein misfolding and aggregation, a key feature for amyloid diseases. Specifically, experimental measurements often consist in the time-dynamics of a moment of the population (*i.e.*, for instance the total polymerised mass, as in Thioflavine T measurements, or the second moment measured by Static Light Scattering).

In a first study [4], in collaboration with H.T. Banks and H. Rezaei, we illustrated the use of tools (asymptotic theories of standard error quantification using appropriate statistical models, bootstrapping, model comparison techniques) in addition to sensitivity that may be employed to determine the information content in data sets. We do this in the context of recent models [87] for nucleated polymerisation in proteins, about which very little is known regarding the underlying mechanisms; thus the methodology we developed may be of great help to experimentalists.

In another study [39], related to a different biological setting (the frog olfactory tract), we use a method based on the Mellin transform, as in [64], to solve a spectral inverse problem arising from the modeling of the transduction of an odor into an electrical signal. The problem is to find the spatial distribution of CNG ion channels along the cilium of a frog, which allow a depolarizing influx of sodium ions, which initiate the electrical signal. This problem comes down to solving a Fredholm integral equation. We prove observability and continuity inequalities by estimating the Mellin transform of the kernel of this integral equation. We perform numerical computations using experimental data.

To get more insight into the estimation of reaction rates and size distributions of protein polymers, we are now developing an approach based on a data assimilation strategy. In this purpose, A. Armiento's Ph.D is focused on setting this framework problem when the experimental measurements consist in the time-dynamics of a moment of the population (*i.e.* for instance the total polymerised mass, as in Thioflavine T measurements, or the second moment measured by Static Light Scattering). In [37] we proposed a general methodology, and we solved the problem theoretically and numerically in the case of a depolymerising system. We then applied our method to experimental data of degrading oligomers, and conclude that smaller aggregates of ovPrP protein should be more stable than larger ones. This has an important biological implication, since it is commonly admitted that small oligomers constitute the most cytotoxic species during prion misfolding process.

7.2.3. Time asymptotics for growth-fragmentation equations

The long-term dynamics of fragmentation and growth-fragmentation equations has been for long an important research field for BANG then MAMBA research team. Thanks to these common efforts, these equations are now well understood. However, there remain some interesting open questions. In particular, if the generic long-time behaviour for the linear equation is known - given by a (generally exponential) trend towards a steady exponential growth described by the positive eigenvector linked to the dominant eigenvalue, see [84] for most recent results - critical cases are not yet fully understood.

With Miguel Escobedo, we focused on an important critical case, when the fragmentation is constant and the growth rate is either null or linear [43]. Using the Mellin transform of the equation, we determine the long time behaviour of the solutions and the speed of convergence, which may be either exponential or at most polynomial according to the subdomain of $(t, x) \in \mathbb{R}_+^2$ which is considered. Our results show in particular the strong dependence of this asymptotic behaviour with respect to the initial data, in contrast to the generic results. Following our study, J. Bertoin and A. Watson proposed a complementary probabilistic analysis of related models [60]. These results exemplify the continuing need for further analysis of these interesting equations.

7.2.4. Cell aggregation by chemotaxis

We follow our investigation on the kinetic model describing the chemotactic motion of bacteria. When taxis dominates the unbiased movements, the kinetic system is approximated by the aggregation equation. The study of such equation is challenging since blow-up in finite-time of solutions occurs. We have defined the notion of measure-valued solution [8] and we have proposed and studied a numerical scheme to simulate these solutions[18].

In another approach, more accuracy can be obtained with the kinetic model by adding an internal variable describing the methylation level of the internal receptors of bacteria. In [55] we have investigated the link between these kinetic models with an internal variable and the one without internal variable.

7.3. Liver modeling

Participants: Noémie Boissier, Dirk Drasdo, Géraldine Cellière, Adrian Friebe, Group Heinzle [Univ. Saarbruecken, Germany], Group Hengstler [IfADo, Germany], Stefan Hoehme, Tim Johann, Irène Reo [Vignon-Clementel], Paul Van Liedekerke, Eric Vibert [Hopital Paul Brousse], Group Zerial [Max-Planck Inst. for Molecular Genetics, Dresden, Germany], Groups Iflow, Notox, Vln.

7.3.1. Ammonia detoxification after drug-induced damage

Overdosing acetaminophen (APAP) is the main reason for acute liver failure in the US and UK. Overdose of APAP destroys the hepatocytes located in the center of each liver lobule (pericentral damage), the repetitive functional and anatomical tissue units of liver. Human has about a million of such lobules. As a consequence, the blood is not sufficiently detoxified from ammonia, which is toxic to the body and can lead to encephalopathy. In France about 1000 cases of ammonia intoxication each year. In recent papers we demonstrated by an integrated model that the widely accepted key reactions scheme of ammonia detoxification is insufficient to explain ammonia detoxification after pericentral lobule damage and predicted a missing ammonia sink [73]. This finding has triggered new experiments leading to the identification of a widely ignored but fundamentally important ammonia sink mechanism We could show by a testing a number of different mechanisms within novel models that this sink mechanism was the only one able to explain the data [15]. The reaction turned out to have the potential to be therapeutically used by injection of a molecular cocktail triggering it. In the animal model death could be prevented using this cocktail hence providing a possible therapy approach for patients suffering from hyperammonemia. [15]. In a follow-up work, further models have been studied and classified by statistical methods to quantify model selection (Cellière et al., in preparation).

7.3.2. Concepts of modeling of liver across all scales: multiscale liver modeling

Based upon developed multiscale concepts [12], we developed a multi-level spatial temporal multiscale models of APAP (paracetamol, acetaminophen) toxicity and ammonia metabolism. In one of these models we integrated molecular pathways of APAP drug toxicity (PD); in another one, we represented the ammonia detoxification pathway into each individual hepatocyte of an agent-based model that describes the precise liver lobule architecture (compare with [73]). This allows us to study the impact of space and architecture on the drug toxicity and drug detoxification. We find in certain cases important differences between models that do represent architecture and those that do not (Cellière et al., in preparation).

7.3.3. Predicting *in vivo* drug toxicity from *in vitro* data

APAP (paracetamol, acetaminophen) *in vitro* experiments have been used to calibrate a model of APAP drug toxicity with *in vitro* data, and modify this model to predict *in vivo* toxicity. This procedure is aimed at as a general pathway among cosmetic and pharmaceutical companies to eliminate or at least reduce animal experiments and it should allow a better prediction of drug toxicity in human. Three critical differences between *in vitro* and *in vivo* settings were stepwise integrated in the model calibrated with *in vitro* toxicity data to study their impact on *in vivo* toxicity predictions: (1) The temporal drug exposure profile, (2) the temporal concentration profile of a class of key enzymes, CYP enzymes. Only in hepatocytes in which CYP enzymes are present, APAP is metabolised and downstream apoptosis can occur. (3) The liver architecture, that is responsible for critical differences in the spatial distribution of the drug. The results are in preparation for publication (Cellière et. al., in preparation).

7.3.4. Miscellaneous

In addition, regenerating lobules after partial hepatectomy were analysed by image analysis, and first simulations of blood and bile flow and molecular transport in those lobules simulated.

7.4. Miscellaneous

Participants: Noémie Boissier, Maria José Cáceres [Universidad de Granada], Julien Chevallier [Université de Nice], Géraldine Cellière, Marie Doumic, Dirk Drasdo, Adrian Friebel, Group Heinzle [Univ. Saarbruecken, Germany], Group Hengstler [IfADo, Germany], Stefan Hoehme, Tim Johann, Group Klingmueller [German Cancer Center, Heidelberg], Johannes Neitsch, Benoît Perthame, Patricia Reynaud [Université de Nice], Group Reo [Inria Paris - Rocquencourt], Paul Van Liedekerke, Eric Vibert [Hopital Paul Brousse], Yi Yin, Group Zerial [Max-Planck Inst. for Molecular Genetics, Dresden, Germany], Groups Iflow, Notox, Vln.

7.4.1. Network formation and neuroscience

Motivated by neurodevelopment and differentiation in developing tissues, a new explanation for sharp boundary formation is analysed in [25]; interestingly, this phenomenon relies on a limited diffusion of homeoproteins (collaboration with the Mycenae team).

Models for neural networks have been proposed which describe the probability to find a neuron for which a time s has elapsed since the last discharge. These are written under the form of a nonlinear age-structured equation where the total network activity modulates the firing rate. An inhomogeneous network of networks with variability on the refractory period is studied in [19].

We have also continued the analysis and numerical simulation of models for natural transportation networks formation based on an elliptic-parabolic system of partial differential equations. The model describes the pressure field using a Darcy's type equation and the dynamics of the conductance network under pressure force effects. Randomness in the material structure is represented by a linear diffusion term and conductance relaxation by an algebraic decay term [16]. Figure 1 below gives a numerical simulation of a network formed by such a model.

7.4.2. Microscopic approach of a time elapsed neural model

The spike trains are the main components of the information processing in the brain. To model spike trains several point processes have been investigated in the literature. More macroscopic approaches have also been studied, using partial differential equation models. With J. Chevallier, M. Cáceres and P. Reynaud-Bouret, we wanted to build a bridge between several point processes models (Poisson, Wold, Hawkes) that have been proved to statistically fit real spike trains data and age-structured partial differential equations as introduced by Pakdaman, Perthame and Salort. To do so, we focused on a seemingly simple one-neuron model, for which we stated the - nonlinear and strongly coupled - PDE model satisfied in average by its point measure when the process model is a Poisson, a Wold or a Hawkes process [10].

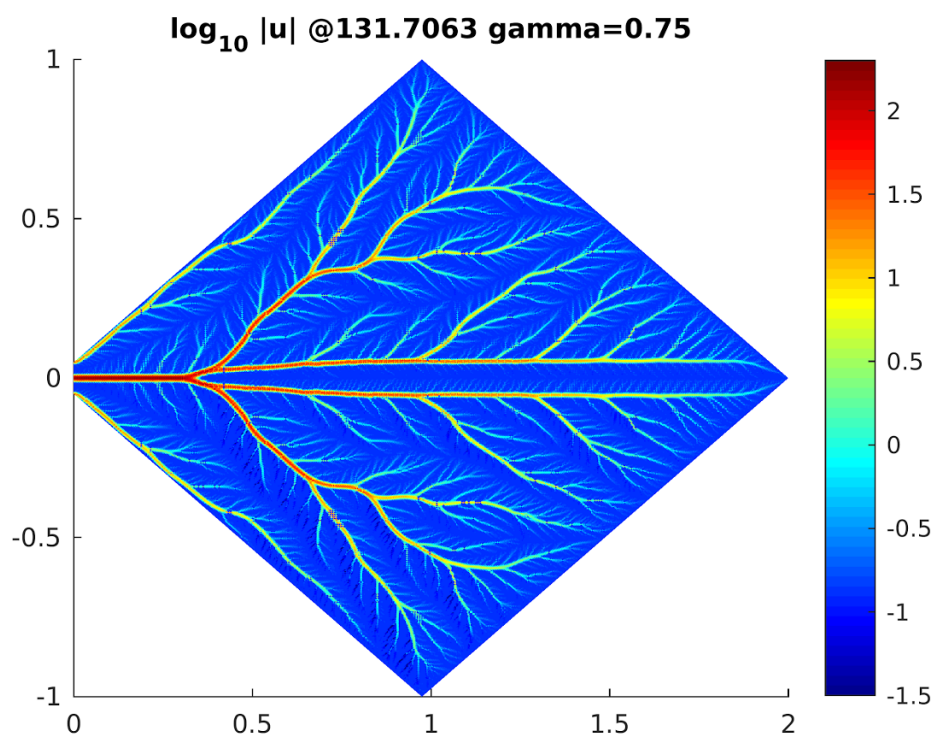


Figure 1. Network formation based on an elliptic-parabolic system of partial differential equations.

7.4.3. Uncertainty propagation

In [42], we study two intrusive methods for uncertainty propagation in scalar conservation laws based on their kinetic formulations. The first one is based on expansions on an orthogonal family of polynomials. The second method uses convolutions based on Jackson kernels. We prove that it satisfies BV bounds and converges to the entropy solution but with a spurious damping phenomenon. Therefore we introduce a second method, which is based on projection on layered Maxellians, and which arises as a minimisation of entropy. This new method satisfies the maximum principle by construction as well as partial entropy inequalities and thus provides an alternative to the standard method of moments which, in general, does not satisfy the maximum principle. Simple numerical simulations for the Burgers equation illustrate these theoretical results.

7.4.4. Simulation of tissue mechanics with agent-based models

In ref. [29] we study and discuss in how far mechanical effects of cells in tissue organisation and growth processes can be captured by agent-based models. We consider a wide range of agent-based models, i.e., lattice base models with one lattice site allowing for many cells or one cell at most, many lattice sites occupied by a single cells (so called Cellular Potts model, Lattice Gas Cellular Automaton approaches, center-based models and vertex models, in which the forces between cells are calculated as forces between the cell centers, as well as deformable cell models in which the cell surface is triangulated. We consider growth of monolayers and multicellular spheroids as reference problems. We also compare in this paper spatial resolution, the capability of the different approaches to represent the physics, cell shape, the computational efficiency and code access. In addition, models evaluating the mechanical effects of growing cell populations in elastic capsules were established and studied.

MIMESIS Team

6. New Results

6.1. Augmented reality for surgery

We have developed a method for real-time augmented reality of internal liver structures during minimally invasive hepatic surgery. Vessels and tumors computed from pre-operative Computed Tomography Angiograms (CTA) scans can be overlaid onto the laparoscopic view for surgery guidance. Compared to current methods, our method is able to locate the in-depth positions of the tumors based on partial three-dimensional liver tissue motion using a real-time biomechanical model. We are pursuing the development of this augmented reality system by using a better biomechanical model, and by relying on parameter optimization and additional per-operative information to further improve accuracy and robustness. In addition, more experiments, and also clinical studies are being performed to precisely measure the benefits and limitations of our approach. This work is strongly related to our involvement in the IHU Strasbourg and is tightly linked to the SOFA-OR project. Many articles were published on this topic [28], [16], [17].

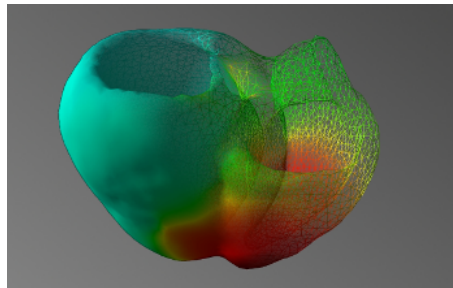


Figure 9. Electrophysiology model of the human heart

6.2. Cardiac electrophysiology

Cardiac arrhythmia is a very frequent pathology that comes from an abnormal electrical activity in the myocardium. This pathology can be treated by catheterization and ablation of the malfunctioning cardiac tissue. The skills required for such interventions are still very challenging to learn, and typically acquired over several years. We first developed a training simulator for interventional electrocardiology and thermoablation of these arrhythmias. Based on physical models 9, this training system reproduces the different steps of the procedure, including endovascular navigation, electrophysiological mapping, pacing and cardiac ablation. Based on a scenario of cardiac arrhythmia, cardiologists assessed the interactivity and the realism of our simulation. This work has been submitted in a journal and is currently under review.

Beyond electrophysiology training, our work around the cardiac electrophysiology also consisted in personalizing our mathematical models. Using the dense electrograms recorded intra-operatively, we presented an accurate and innovative approach to personalize our model, i.e. estimate patient-specific parameters. The modeling in silico of a patient electrophysiology is needed to better understand the mechanism of cardiac arrhythmia.

6.3. Cryoablation

In 2015, we carried on the work around thermal ablation and pre-operative planning based on a thermal Finite Element Model (FEM). The cryoablation technique consists in inserting needles that freeze the surrounding tissues, thus immediately leading to cellular death of the tissues. Cryoablation procedure is used in many medical fields for tumor ablation, and even starts being used in cardiology. In this scope, we built a simulator **10** able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues.

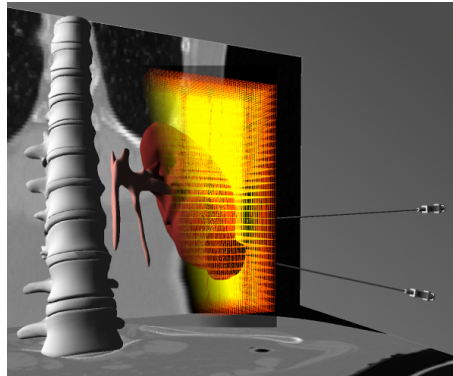


Figure 10. Cryosurgery simulation with the creation of an iceball in the kidney

6.4. Lipofilling reconstructive surgery

We have developed a method to simulate the outcome of reconstructive facial surgery based on fat-filling. Facial anatomy is complex: the fat is constrained between layers of tissues which behave as walls along the face; in addition, connective tissues that are present between these different layers also influence the fat-filling procedure. To simulate the end result, we have proposed a method which couples a 2.5D Eulerian fluid model for the fat and a finite element model for the soft tissues. Both models are coupled using the computation of the mechanical compliance matrix. We had two contributions: a solver for fluids which couples properties of solid tissues and fluid pressure, and an application of this solver to fat-filling surgery procedure simulation. Vincent Majorczyk defended his PhD [14] on this topic in 2015.

6.5. Neurosurgery

Based on an intra-operative registration method, we developed a simulation of a DBS (Deep Brain Stimulation) surgery which can help the surgeon to locate anatomical structures for a safer and a more efficient treatment. The method relies on the biomechanical model of brain shift we developed during the last years. Because some parameters of the model are unknown, we propose to estimate them with an optimization process. The cost function evaluates the distance between the model and the segmentation of pneumocephalus, the only indicator of brain shift visible on an intra-operative CT scan. In 2015, an article about the rest shape of the brain was accepted [19].

6.6. Physics-based registration algorithms

Before targeting the augmented reality for laparoscopic operations, an important step consists in solving the initial alignment problem. Given a pre-operative image of the organ (usually a CT scan) a detailed mesh is constructed. To make the information stored in this mesh available during the operation, the mesh must be registered onto the intra operative view. However, mainly due to the pneumoperitoneum, the organ has

undergone important deformation between the pre-operative images acquisition and the operation. The pre-operative shape and the intra-operative shape of the organ do not correspond. Therefore a non rigid registration is required to align the mesh and the real organ. Our registration algorithms also allowed us to work on means to automatically recover boundary conditions of a patient specific liver.

We created a statistical atlas of the human liver to store the positions of the boundary conditions: the veina cava and the anchor point of the falciform ligament positions. This method was accepted at ABME in 2015 [21]. We also developed a new registration method that evolves automatically from a rigid registration to a non rigid registration to solve the initial alignment problem. The method uses some anatomical features of the liver such as the anchor point position of the falciform ligament.



Figure 11. Detecting a catheter in interventional medical images

6.7. Radiation-less guidance during interventional radiology procedures

Significant changes have taken place over the past 20 years in medicine with the development of minimally-invasive procedures. While surgery evolved towards laparoscopy for instance, interventional radiology has become another alternative for many pathologies. Yet, some limitations remain: for percutaneous procedures, soft tissue motion, either due to breathing or deformation induced by the needle, changes the location of the target. When using image guidance, or robotic control, this remains a major obstacle. Regarding catheter-based interventions, the lack of 3D information, and extensive use of X-ray imaging to visualize the path to be followed, are among the main issues. We propose to address these different problems by developing an advanced navigation system which relies on a combination of real-time simulation and information extracted from intra-operative images to assess the current position of the needle. Such a method would have direct applications in pre-operative planning, per-operative guidance, and control for robotics. Our approach will combine advanced modeling of the device, soft tissue deformation, tissue-tool interactions, and planning algorithms 11 .

6.8. Regional anaesthesia

The RASimAs project (Regional Anaesthesia Simulator and Assistant) is a European research project funded by the European Union's 7th Framework Program. It aims at providing a virtual reality simulator and assistant to doctors performing regional anaesthesia by developing the patient-specific Virtual Physiological Human models. In this project, we are in charge of developing a simulation of a needle inserted into a leg using the SOFA framework 12 . We especially focused on the integration of the needle simulation into SOFA. We planned to release the first version of the simulator by January 2016.

In the context of RASimAs, we organized a coding sprint in Strasbourg in April 2015.

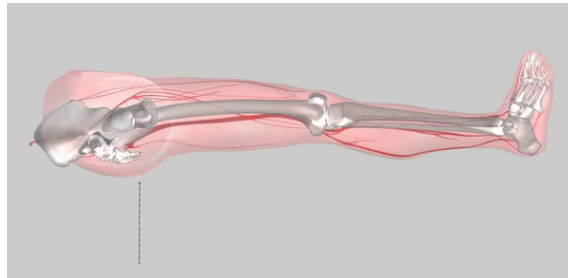


Figure 12. Needle insertion in a muscle in the context of local anaesthesia

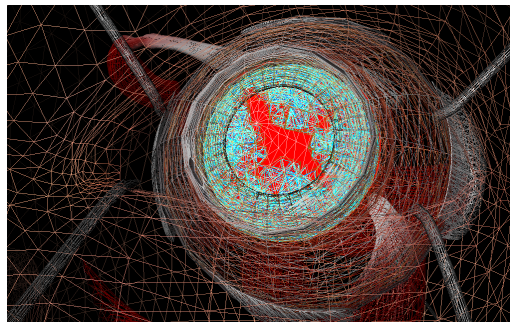


Figure 13. FEM model of the eye used in our simulation of retina surgery

6.9. Training for retina surgery

Retina surgery is an increasingly performed procedure for the treatment of a wide spectrum of retinal pathologies. Yet, as most micro-surgical techniques, it requires long training periods before being mastered. To properly answer requests from clinicians for highly realistic training on one hand, and new requirements for accreditation or recertification from surgical societies on the other hand, we are developing a high-fidelity training system for retinal surgery. This simulator will be built upon our strong scientific expertise in the field of real-time simulation, and a success story for technology transfer in the field of cataract surgery simulation. Members of the consortium have a long expertise in the development of prototypes, as well as collaborations with clinical partners. The simulation system that we propose to develop is based on the Open Source simulation platform SOFA, and relies on expertise from our partners to ensure clinical and industrial relevance. This work is initially funded through the ANR project RESET which started in March 2015. A first version [13](#) of the training system has been delivered and we made a live demonstration at the Journée Alsacienne d’Ophtalmologie.

6.10. Virtual Cutting

The simulation of cutting is a central interest in the team. We especially work on the simulation of surgical cuts [14](#), tearing and other separations of materials induced by surgical tools. On the one hand, we investigated the theoretical aspect: using the standard finite element method (FEM) combined with a re-meshing approach, we replace locally the current structure of the mesh in order to allow for a separation. On the other hand, we detected a separation in the motion of an object provided by a monocular video stream. With that detection, we can provide an augmented reality during the cutting and tearing of a deformable object.

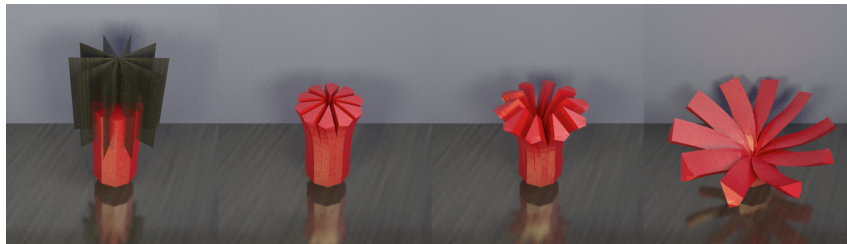


Figure 14. Our cutting algorithm in SOFA

The theoretical aspect of our work has been published in an article both at the conference Computer Graphics International CGI [\[24\]](#) and in the journal "The visual computer" [\[20\]](#). The application in augmented reality [15](#) has been published at two conferences: "Augmented Reality during Cutting and Tearing of Deformable Objects", International Symposium on Mixed and Augmented Reality (ISMAR) [\[30\]](#).

To read more about our projects and results, please visit our website: <http://mimesis.inria.fr>.

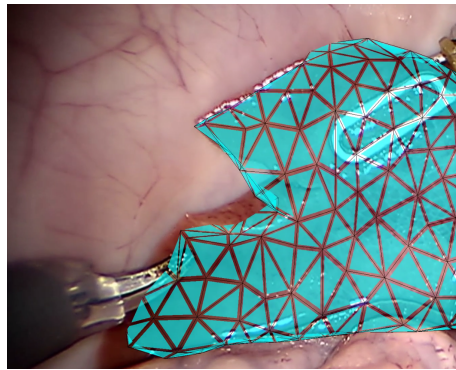


Figure 15. Augmented reality on a liver involving large deformation and cutting, i.e. topological changes

MNEMOSYNE Project-Team

7. New Results

7.1. Overview

Though our view is systemic, our daily research activities are also concerned with the design, at a given scale of description, of models of neuronal structures, each concerned with a specific learning paradigm. Of course, a major challenge is to integrate these elements in a systemic view, i.e. to put a specific emphasis on the way each neuronal structure communicates with the rest of the system and to highlight how its learning paradigms interact with other memory systems.

Among the numerous loops involving the brain, the body and the environment, a basic grid of description corresponds to distinguish “perception aspects of loops”, the goal of which is to extract from the inner and outer world sensory invariants helpful to identify and evaluate the current state and to make predictions from previous learning, and “action aspects of loops”, the goal of which is to rely on this sensory and emotional information to decide, plan and trigger actions for the benefit of the body.

This year, our team was engaged on the following topics: Concerning perception aspects of loops, we published original models of the amygdala and of the hippocampus and considered their role in pavlovian conditioning and their evaluation as classical models in machine learning. Concerning action aspects of loops, in addition to a critical analysis of the current views of the interactions between the prefrontal cortex and the basal ganglia [15], we have proposed an original model for the formation of habits and have also studied related theoretical problems in machine learning, for data representation. Finally, we also report here more methodological achievements, corresponding to the design of algorithmic ersatz of cerebral subsystems.

7.2. Pavlovian conditioning

Within perception aspects of loops, pavlovian conditioning is a very interesting learning paradigm to study in a systemic view because it is tightly related to other learning paradigms like episodic and semantic memory. This year, we have published papers presenting the biological basis of models of two fundamental structures in pavlovian conditioning, the amygdala [1] and the hippocampus [2]. We have also evaluated their most critical features, when considered as models of machine learning, namely their architecture and implementations at both rate and spiking levels [10] and their robustness to interference [11].

7.3. The formation of habits

Concerning action aspects of loops, we made important extensions to a model of basal ganglia that we developed recently [54] in interaction with another team in our neuroscience lab. In addition to extending the bio-plausibility of this model with an implementation at the spiking level, we have also developed this year a new theoretical framework that provides a novel explanation for the formation and the expression of habits in the cortex of primates by considering the basal ganglia as an implicit supervisor. This has been achieved with a model of basal ganglia running both at the rate and spiking levels. This framework predicts that Hebbian learning and reinforcement learning can be explicitly dissociated by inactivating the output of the basal ganglia during learning and later tested in normal conditions. Experimental results in the monkey confirmed this prediction and explain how a behavioral decision results from both the cooperation (acquisition) and competition (expression) of two distinct but entangled memory systems.

7.4. Beyond symbolic models

Using a biologically plausible model, we have been investigating some external and internal factors related to the stimulus representation that might affect the decision making and action selection [41]. We used a computational model of the cerebral structure Basal Ganglia, inspired and replicated from a model designed in previous studies [54]. One of the questions we attempt to address is to what extent the physical properties of the stimulus affect the decision to overcome the impact of reward associated to the stimuli.

7.5. Algorithmic ersatz of cerebral subsystems

As far as the systemic modeling and simulation of high-level brain functions are concerned (e.g., sensory-motor behavior, action selection and planning, perceptual categorization), we need to confront biologically plausible models at different scales of description with functional models that are not constrained by biological facts but still reproduce the expected functional response. This is mandatory to benchmark bio-physical models with respect to their equivalent in classical machine learning, in order to evaluate the degree of naiveness of their performances and also to build feasible simulation in which detailed biological models can interact with less plausible modules in order to be evaluated in realistic numerical situations.

This year, a set of formalism such as the Friston free-energy minimization general principle, deep-learning and related architectures, and more specific formalisms such as harmonic control or adaptive-subspace self-organizing maps have been studied and reviewed. The next step is to write a review, with the challenge of proposing an unifying view of those, and at a more concrete level, to propose the integration of a relevant subset of the related algorithms as a easily usable toolbox. This can be particularly useful to design global models of cognitive functions, even if biologically-inspired models are not yet available for all their components.

Preliminary key points regarding numerical experimentations have been published in this methodological paper [16].

MODEMIC Project-Team

7. New Results

7.1. Mathematical models for microbial ecology

7.1.1. Differential equations models

Participants: Jérôme Harmand, Claude Lobry, Alain Rapaport, Yessmine Daoud, Sonia Hassam, Zeyneb Khedim, Alejandro Maximiliano Rojas.

Anaerobic digestion refers to the transformation of biodegradable material by micro-organisms in absence of oxygen (it can be found in waste-water treatments or industrial fermentation, and occurs naturally in soils). It receives an increasing consideration due to recent technological advances, but also because it is a source of renewable energy (bio-gas, fuel...). The anaerobic digestion is a complex set of bio-processes, for which there is a strong expectation of tractable models. We have proposed and studied new mathematical models that takes into account the following features:

- Microbial food chains are present in anaerobic digestion where the different reaction steps can be seen as such: the waste products of the organisms at one trophic level (i.e. one reaction step) are consumed by organisms at the next trophic level (i.e. the next reaction step). In [54] we study a model of a two-tiered microbial ‘food chain’ with feedback inhibition, which was recently presented as a reduced and simplified version of the anaerobic digestion model ADM1 of the International Water Association (IWA) (cf. [61]). It is known that in the absence of maintenance (or decay) the microbial ‘food chain’ is stable. In [61], using a purely numerical approach and ADM1 consensus parameter values, it was shown that the model remains stable when decay terms are added. In [54] we prove that introducing decay in the model preserves stability whatever its parameters values are and for a wide range of kinetics.
- In the thesis by Sonia Hassam [13], we have proposed a new procedure to easily and systematically obtain a simple model useful for control purposes of any process for which an ADM1 is available. The simplified model has two major characteristics : its states keep their physical meaning and it remains nonlinear. The technique is based on the state association technique proposed in [26].
- Zeyneb Khedim (University of Tlemcen, Algeria) has began her PhD in 2014. She is working on the modeling and control of anaerobic digestors. In particular, she works on the reduction of models using the state association approach proposed by Sonia Hassam but for substrates highly loaded in nitrogen such as algae. She has published this year a survey with Sonia Hassam [36].
- Yessmine Daoud (ENIT-LAMSIN, Tunis, Tunisia) continues her work on the analysis of a model of the literature to optimize anaerobic processes [35]. She is preparing a journal paper which should be submitted during 2016.

Formerly, the team has studied chemostat models where the bacterial compartment is split into “planktonic” and “attached” bacteria (such as in flocculation or biofilms formation), under the hypothesis that attachment and detachment are fast phenomena. Under certain mixing conditions, this last condition is no longer satisfied. We have studied on the non-reduced model the competition between a species that presents growth inhibition in planktonic form with a species that does not attach. This consideration leads to multiple positive equilibria but surprisingly it can also conduct to limit cycles [53] (paper under revision for Applied Math. Model.).

Spatial heterogeneity is often observed in non perfectly mixed bioprocesses or in populations in natural environments. The representation of spatial heterogeneity in population models with patches or interconnected models, rather than p.d.e., is one of the specialties of the team, that allows us to characterize non intuitive effects of spatialization :

- The very basic Rosenzweig-MacArthur model is subject to the "atto-fox" problem [2] when considered for homogeneous populations. Is it still true in case of heterogeneous populations? The idea is: the resource population being not small at the same time in different places is it possible that, thanks to dispersal, it will not disappear? One possible idealization of heterogeneous populations is to use reaction-diffusion equations. We do not take this direction for two reasons
- (i) Due to the presence of a limit cycle in the homogeneous system mathematics of such reaction diffusion are difficult.
- (ii) Idealization through reaction-diffusion is not the best one; patch-systems (or lattice differential equations in mathematical terms) are better in many cases.

Our ultimate objective is to provide mathematical results for systems with a large number of patches but, as a first step, in the paper [27] we consider two patches. It is proved that for some migration rates, stable periodic solutions avoiding "atto-fox" exist.

- The standard model for the dynamics of a fragmented density-dependent population is built from several local logistic models coupled by migrations.

First introduced in the 1970s and used in innumerable articles, this standard model applied to a two-patch situation has never been completely analyzed. The motivation for studying this problem came out from discussions at the Bernoulli semester organized in 2014 and 2015 by the team at the EPFL (see the 2014 activity report and Section 8.3.3.1). It addresses very fundamental issues in theoretical ecology. In the paper [15] written in collaboration with R. Arditi (U. Fribourg) and T. Sari (IRSTEA Montpellier), we complete this analysis and we delineate the conditions under which fragmentation associated to dispersal is either beneficial or detrimental to total population abundance. Therefore, this is a contribution to the SLOSS question. Importantly, we also show that, depending on the underlying mechanism, there is no unique way to generalize the logistic model to a patchy situation. In many cases, the standard model is not the correct generalization. We analyze several alternative models and compare their predictions. Finally, we emphasize the shortcomings of the logistic model when written in the r-K parameterization and we explain why Verhulst's original polynomial expression is to be preferred.

- We have carried on our former work on the role of particular interconnections patterns on the global stability of chemostat model with inhibition. While we focused formerly on the conditions for which a spatial structure ensures the global stability when the chemostat model is bi-stable in homogeneous environment, we have shown that at the opposite a spatial structure can make unstable the dynamics of the chemostat model with inhibition when it is stable in a homogeneous environment [30].
- In collaboration with Géosciences Rennes (Jean-Raynald de Dreuzy, Tristan Babey) and in the scope of the co-supervision of the PhD of Alejandro Rojas (also in the collaboration within the associated team with Chile), we have carried on the complete equivalence between several models used in Geosciences to characterize soil fractures : MINC (Multiple INTERacting Continua), MRMT (Multi-Rate Mass Transfer) and SINC (Structured INTERacting Continua). We have shown that the irreducibility of the network graph is not sufficient to obtain equivalence : a controllability assumption has also to be fulfilled [42]. Moreover, this kind of models has been used to fit experimental data of reconstituted soils at Inra Grignon and has shown the role of convection in the acquisition of pesticides by micro-organisms [46] (paper in preparation). This work will be continued in the framework of the new ANR project Soil μ 3D (see Section 8.2.1).

In resources/consumers models, heterogeneity can be also due to time varying inputs of resources (e.g. light in micro-algae populations). While, most of the literature studies periodic inputs, we have begun investigations of more general time varying inputs in chemostat like models, having in mind to characterize "pull-back attractors" (rather than forward attractors) [43].

7.1.2. Stochastic and hybrid discrete-continuous dynamical models

Participants: Bertrand Cloez, Claude Lobry.

7.1.3. Approximation of quasi-stationary distributions

The study of the long-time behavior of a stochastic process is one of the main questions of interest for modeling. In a standard Markov setting, this leads to the study of the convergence towards the invariant distribution. However, in many applications such as population dynamics for instance, the stochastic dynamics is killed in a finite (random) time so that the standard asymptotic regime is trivial. In this case, it can be interesting to focus on the behavior of the process conditionally to its non-extinction before a given time t . Under appropriate assumptions, one can exhibit a convergence of this conditional distribution towards a law called *Quasi-Stationary Distribution*. Properties of this law is then fundamental. In [21], we study an algorithm to approximate this distribution and we provide proof of convergence as well as precise rates for convergence. This one is based on a reinforced random walk.

7.1.3.1. Lotka Volterra in fluctuating environment

In the paper [49], we consider two dimensional Lotka-Volterra systems in a fluctuating environment. Relying on recent results on stochastic persistence and piecewise deterministic Markov processes, we show that random switching between two environments that are both favorable to the same species can lead to the extinction of this species or coexistence of the two competing species. This work has been accepted in Journal of applied probabilities, provided major revisions. We submitted a new version with the new title: Lotka Volterra with randomly fluctuating environments or "how switching between beneficial environments can make survival harder".

7.2. Analysis and supervision of bioprocesses

7.2.1. Models development and identification

Participants: Yessmine Daoud, Jérôme Harmand, Nesrine Kalboussi, Guilherme Pimentel, Alain Rapaport.

Membrane bioreactors combine a filtration process (with a membrane) and a suspended growth rate bioreactor. This recent technology present many advantages compared to conventional ones, but is more sophisticated and requires refined control because of possible problems related to the risk of membrane fouling. After the PhD by Amine Charfi defended in 2014 we continue to work on the modeling and control of membrane bioreactors.

- Within this framework, new results have been obtained and a new model including several fouling mechanisms has been proposed [22].
- In the scope of the PhD of Guilherme Pimentel (defended in February 2015, [11]), we have proposed a simple three time scales model in view of the control of the cake formation [14], [33]. This model has been validated on real data from a pilot plant at Univ. Mons (Belgium).
- The PhD thesis by Nesrine Kalboussi (ENIT-LAMSIN, Tunis, Tunisia) has just begun. It is dedicated to the early detection and control of membrane fouling. At present time, Nesrine is working on the bibliography about modeling and control of membrane bioreactors.

In many bioprocesses models, the loss of nutrient used for the maintenance of bacteria is neglected compared to the important nutrient supply. In poor environment, such as natural one in the oceans, this is no longer verified. In our collaboration with the LOMIC lab (Banyuls-sur-Mer), we have shown that the consideration of a maintenance term in the chemostat model allows to fit the data observed in experimental chemostats, and moreover that the level of maintenance is correlated to the activities of bacteria under the presence of light [23]. This gives a possible explanation of the variable yield observed in the bacterial compartment of marine ecosystems.

7.2.2. Synthesis of control laws

Participants: Térrence Bayen, Walid Bouhafs, Amel Ghouali, Jérôme Harmand, Claude Lobry, Guilherme Pimentel, Alain Rapaport, Victor Riqueleme.

We investigate two kinds of bioprocesses to be controlled, arising in industrial biotechnology (digesters, wastewater purification...) or in the bioremediation of natural environments (lakes, landfill...).

7.2.2.1. Industrial biotechnology

In the framework of the PhD of Guilherme Pimentel [11] and the pilot plant at Mons, a nonlinear predictive law based on the model exposed in 7.2.1 has been tested and validated for piloting the process[34].

Control of biological reactors are still of great interest, notably but not only with respect to anaerobic digestors that can be destabilized due to the accumulation of intermediate metabolites that can inhibit the growth of some bacteria.

- Amel Ghouali (Cotutelle Univ. Montpellier and Univ. of Tlemcen, Algeria) who has defended her PhD in December has developed an optimal control strategy to optimize the production of biogas over a given period of time [12]. In particular, she has solved an original optimal control problem using the maximum principle of Pontryagin [25].
- Within the scope of the PhD thesis by Walid Bouhafs (ENIT-LAMSIN, Tunis, Tunisia), we have proposed a new optimal control strategy for systems in which two specific substrates are degraded by two different bacterial consortia, one being limited by the oxygen while the other is inhibited. Walid will defend his PhD in next February.

The minimal time criterion is of particular interest in biotechnology, as it leads to time-independent feedback controllers.

- The paper [19] is devoted to the study of the minimal time problem of a fed-batch reactor, under the presence of a saturation point on the singular locus (this typically occurs whenever the growth rate function is of Haldane type and when typically the maximum input flow rate is not high enough to maintain the substrate concentration constant). This brings non-intuitive issues for the optimal synthesis (existence of switching curve and point of prior saturation).
- In the work [47], we study the minimal time control to drive a chemostat model to a target point. Such a problem finds application typically in the case where the input substrate concentration changes yielding in a new steady state. Converging fast towards this new reference point is much desired in practice. One essential feature of the present work is that the system takes into account a recirculation of biomass (as it is more and more often the case in modern biotechnology). We depict an optimal synthesis and provide an optimal feedback control by using the Pontryagin Maximum Principle and geometric control theory for both Monod and Haldane kinetics.

7.2.2.2. Bioremediation of natural environments

In the scope of the associated team with Chile (see 8.3.1.1) and the co-supervision of the PhD of Victor Riquelme, we have carried on the study of optimal syntheses for the minimal time treatment of natural water reservoirs (such as lakes) [41]. We have proved that the minimal time strategy consists in a most-rapid approach to homogeneous concentrations, even though the optimal control problem is non convex. Moreover, we have shown that a large diffusion increases the treatment time when the resource is everywhere highly polluted, while it can at the opposite be beneficial when only part of the resource is polluted (paper under revision for SIAM J. Cont. & Optim.). This feature should serve the practitioners in the choice of pumps positioning in a originally clean water resource that is suddenly affected by a local pollution. Moreover, we have shown, in collaboration with A. Rousseau (EPI LEMON), how these analytic feedback laws obtained on a over-simplified representation of the spatial heterogeneity behave quite satisfactorily when simulated [17]. This year we have started to study to problem of treating two different pollutants, with a anaerobic/aerobic process in series.

Also in the scope of the associated team with Chile, we have characterized the optimal strategy to treat as fast as possible a landfill with the recirculation as a manipulated variable [40], [29], based on a model that we have proposed last year. In presence of singular arcs that are non-admissible (in the sense that the upper bound of the recirculation pump does not allow to stay on the singular arc), we have shown that a kind of *anticipation* law is necessary before operating optimally the switching. This analyses reveals several sub-domains for which the optimal policy requires different kind of measurements. Knowing in which sub-domain the initial stage

of landfill could be inform then the practitioners about which concentration (leachate or solubilized or both) should be ideally measured. This primarily work has led to the co-development of a software mock-up with Chilean partner (see 6.1), in order the study the consideration of spatial heterogeneity in landfill, with the approach exposed in 7.1.1 .

This year, again in the scope of the associated team with Chile and Inria Chile, we have begun a new investigation on modeling and control strategy for the regulation of a lagoon that communicates temporarily with the sea and whose water is exploited by pumping.

7.2.2.3. Theoretical development

The time crisis is an interesting criterion that measures the time spent by a system in a “bad” zone or in “danger”. Typically, when a desired species is under a given (low) threshold, one can consider that this defines a crisis domain. For controlled system, the minimal time crisis has already been proposed in the literature [56]. Nevertheless, only sufficient conditions (i.e. characterization of the solutions of the associated Hamilton-Bellman-Jacobi equation) have been given, and no necessary conditions have been yet proposed, due to the lack of continuity of the integrand cost. We have proposed a regularization of this problem by a family of optimal control problems for which the usual necessary conditions can be derived, and studied the convergence [20]. Practically, this allows to use classical software, such as Bocop, to approximate the optimal solutions. In the internship of C. Romero (U. Chile), this technique has been successfully applied on the Lotka-Volterra model with a control on the predator, and a threshold on the prey.

Monc Team

7. New Results

7.1. Axis 1: Tumor modeling for patient-specific simulations

7.1.1. Lung metastasis

Patient specific simulation of tumor growth, response to the treatment and relapse of a lung metastasis: a clinical case [10], [1]

Team participants: Thierry Colin, Julien Jouganous, François Cornelis (Hôpital Pellegrin), Olivier Saut

Other participant: Jean Palussière (Bergonié Institute)

In this work, a parametrization strategy based on reduced order methods is presented for tumor growth PDE models. This is applied to a new simple spatial model for lung metastasis including angiogenesis. The goal is to help clinicians monitoring tumors and eventually predicting its evolution or response to a particular kind of treatment. To illustrate the whole approach, a clinical case including the natural history of the lesion, the response to a chemotherapy and the relapse before a radiofrequency ablation is presented.

Nenuphar

Team participants: Thierry Colin, Julien Jouganous, Marie Martin, Olivier Saut

This work concerns the development of *Nenuphar* which is a software devoting to the evaluation and the surveillance of the tumor aggressiveness.

7.1.2. Take into account the drug resistance

Modeling and analysis of tumor heterogeneity during treatments resistance: GIST liver metastases case

Team participants: Thierry Colin, François Cornelis, Guillaume Lefebvre, Clair Poignard, Olivier Saut

This work deals with tumor heterogeneity analysis and modeling during treatments resistances. A patient-dependent PDEs model, that takes into account two kinds of treatments, is presented. It qualitatively and quantitatively reproduces the different stage during the tumor growth undergoing treatments. In order to overcome a numerical instability linked to the type of modeling, a new numerical scheme is built. Then, an image synthesis method is developed to enable a better comparison between the numerical results and the clinical data. Finally, a robust criteria that quantifies the tumor heterogeneity from the clinical data and from the synthesis images, is built.

Mathematical study and asymptotic analysis of a model for tumour drug resistance [19]

Team participants: Thierry Colin, Thomas Michel, Clair Poignard

In this work we study a partial differential equations model for tumour growth taking into account drug resistance. It is well known that angiogenesis, the process of creation of new blood vessels from existing ones, is induced by tumour cells to get the amount of nutrients and oxygen needed to continue their proliferation when the tumour has reached a critical size. Angiogenesis is therefore a target for therapy. The model we study takes into account two kinds of treatments: a cytotoxic treatment and a treatment which is both cytotoxic and anti-angiogenic. It is based on mass-balance equations on cells densities coupled with a diffusion equation for the nutrients and oxygen concentration. In a first part we prove that the model is well-posed if the initial tumour is compactly supported in the domain, which is the case for tumour metastases. The proof states that the tumour remains compactly supported in a finite time. In the model, we also consider the presence of a necrotic compartment composed of dead cells. Since some tumours can present necrosis while other do not, we want a model which can reproduce these two different cases. The second part of this work is devoted to an asymptotic analysis which proves that the absence of necrosis is the limit case of our model when the necrosis is immediately evacuated.

7.1.3. Motility phenotype

TMOD-03 * Motility controls growth and progression patterns of glioblastoma multiforme [13]

Team participants: Olivier Saut, Thierry Colin

Other participants: Hassan Fathallah, Elizabeth Scribner

Purpose: Glioblastoma multiforme (GBM) is a malignant brain tumor with poor prognosis and high morbidity due to its invasiveness. Hypoxia-driven motility (HM) and concentration-driven motility (CM) are two mechanisms of GBM invasion in the brain. The use of anti-angiogenic drugs has uncovered new progression patterns of GBM associated with significant differences in overall survival times. Here, we test the hypotheses that the types and rates of GBM motility predict its progression pattern and the patients' survival times. **Methods:** We applied a mathematical model of GBM growth and invasion in humans to simulate a clinical trial and study the effects of the rate and mechanism of motility on the patterns of progression and on survival times. **Results:** The motility phenotype appears to determine the progression pattern as well as the survival time of a patient treated by anti-angiogenesis. Highly-dispersive tumors are associated with the longest survival times ($p < 0.001$) and with progression by Expanding FLAIR. Moderately-Dispersive tumors are associated with short survival times and with progression by Expanding FLAIR + Necrosis. Tumors with HM are associated with the shortest survival times and with progression by Expanding Necrosis. The survival times of the latter are similar to non-responders. This investigation also uncovered the HM-CM principle: the aggressive HM-dependent phenotype surfaces only when the rate of CM is low in both untreated and bevacizumab-treated GBM. **Conclusions:** Finding that the motility phenotype is a fundamental property that controls progression and survival times, has biological, clinical and therapeutic implications.

7.2. Axis 2: Bio-physical modeling for personalized therapies

7.2.1. Electroporation

Non-Linear Steady-State Electrical Current Modeling for the Electroporation of Biological Tissue [8]

Team participants: Clair Poignard, Michael Leguebe

Other participants: Marie Breton, Lluís M. Mir (Vectorology and Anticancer Therapies), Francois Buret, Riccardo Scorretti, Damien Voyer, Laurent Krähenbühl (Ampère Laboratory (Lyon) participants), Ronan Perrussel (LAPLACE - Laboratoire Plasma et Conversion d'Énergie, Toulouse)

We propose a non-linear steady-state model of irreversible electroporation in a biological tissue. The non-linear problem is solved using a modified fixed point iteration. The unknown parameters are experimentally estimated from the observation of the necrosis on a potato tissue for different applied voltages. A variability study of the parameters involved in the model is performed.

A second-order Cartesian method for the simulation of electroporation cell models [12]

Team participants: Clair Poignard, Michael Leguebe

Other participant: Lizl Weynans (Memphis team, Inria)

In this work, we present a new finite differences method to simulate electropermeabilization models, like the model of Neu and Krassowska or the recent model of Kavian et al. These models are based on the evolution of the electric potential in a cell embedded in a conducting medium. The main feature lies in the transmission of the voltage potential across the cell membrane: the jump of the potential is proportional to the normal flux thanks to the well-known Kirchoff law. An adapted scheme is thus necessary to accurately simulate the voltage potential in the whole cell, notably at the membrane separating the cell from the outer medium. We present a second-order finite differences scheme in the spirit of the method introduced by Cisternino and Weynans for elliptic problems with immersed interfaces. This is a Cartesian grid method based on the accurate discretization of the fluxes at the interface, through the use of additional interface unknowns. The main novelty of our present work lies in the fact that the jump of the potential is proportional to the flux, and therefore is not explicitly known. The original use of interface unknowns makes it possible to discretize the transmission conditions with enough accuracy to obtain a second-order spatial convergence. We prove the second-order spatial convergence in the stationary linear one-dimensional case, and the first-order temporal convergence for the dynamical non-linear model in one dimension. We then perform numerical experiments in two dimensions that corroborate these results.

Cell membrane permeabilization by 12-ns electric pulses: Not a purely dielectric, but a charge-dependent phenomenon [15]

Team participants: Clair Poignard, Michael Leguèbe

Other participants: Aude Silve (KIT - Karlsruhe Institute of Technology), Isabelle Leray, Lluís M. Mir (Université Paris Sud)

Electric pulses of a few nanoseconds in duration can induce reversible permeabilization of cell membrane and cell death. Whether these effects are caused by ionic or purely dielectric phenomena is still discussed. We address this question by studying the impact of conductivity of the pulsing buffer on the effect of pulses of 12 ns and 3.2 MV/m on the DC-3F mammalian cell line. When pulses were applied in a high-conductivity medium (1.5 S/m), cells experienced both reversible electropermeabilization and cell death. On the contrary, no effect was observed in the low-conductivity medium (0.1 S/m). Possible artifacts due to differences in viscosity, temperature increase or electrochemical reactions were excluded. The influence of conductivity reported here suggests that charges still play a role, even for 12-ns pulses. All theoretical models agree with this experimental observation, since all suggest that only high-conductivity medium can induce a transmembrane voltage high enough to induce pore creation, in turn. However, most models fail to describe why pulse accumulation is experimentally required to observe biological effects. They mostly show no increase of permeabilization with accumulation of pulses. Currently, only one model properly describes pulse accumulation by modeling diffusion of the altered membrane regions.

7.2.2. Cell protrusion

Free boundary problem for cell protrusion formations: theoretical and numerical aspects [20]

Team participants: Olivier Gallinato, Clair Poignard

Other participants: Masahito Ohta (Tokyo University of Sciences), Takashi Suzuki (Osaka University)

In this work, we derive a free boundary problem for cell protrusion formation in which the cell membrane is precisely described thanks to a level-set function, whose motion is due to specific signalling pathways. The model consists in Laplace equation with Dirichlet condition inside the cell coupled to Laplace equation with Neumann condition in the outer domain. The motion of the interface is due the gradient of the inner quantity. We prove the well-posedness of our free boundary problem under a sign condition on the datum similarly to the Taylor criterion in water waves. We also propose an accurate numerical scheme to solve the problem and we exhibit the main biological features that can be accounted for by the model. Even though simplistic from the modeling point of view, we claim that this work provides the theoretical and numerical grounds for single cell migration modeling. In particular, specific chemical reactions that occurred at the cell membrane could be precisely described in forthcoming works.

7.3. Axis 3: Quantitative cancer modeling for biological and preclinical studies

7.3.1. Modelling of metastasis development

Computational Modelling of Metastasis Development in Renal Cell Carcinoma [2]

Team participants: Etienne Baratchart, Sébastien Benzekry, Thierry Colin, Olivier Saut

Other participants: Andreas Bikfalvi, Lindsay S. Cooley, Raphaël Pineau, Wilfried Souleyreau (LAMC - Laboratoire Angiogenèse et Micro-environnement des Cancers), Emeline J Ribot (RMSB - Résonance magnétique des systèmes biologiques)

To improve our understanding of the biology of the metastatic colonization process, we conducted a modelling study based on multi-modal data from an orthotopic murine experimental system of metastatic renal cell carcinoma. The standard theory of metastatic colonization usually assumes that secondary tumours, once established at a distant site, grow independently from each other and from the primary tumour. Using a mathematical model describing the metastatic population dynamics under this assumption, we challenged the theory against our data that included: 1) dynamics of primary tumour cells in the kidney and metastatic cells in the lungs, retrieved by green fluorescent protein tracking, and 2) magnetic resonance images (MRI) informing on the number and size of macroscopic lesions. While the model could fit the primary tumour and total metastatic burden, the predicted size distribution was not in agreement with the MRI observations. Moreover, the model was incompatible with the growth rates of individual metastatic tumours. To explain the observed metastatic patterns, we hypothesised that metastatic foci derived from one or a few cells could aggregate, resulting in a similar total mass but a smaller number of metastases. This was indeed observed in our data and led us to investigate the effect of spatial interactions on the dynamics of the global metastatic burden. We derived a novel mathematical model for spatial tumour growth, where the intra-tumour increase in pressure is responsible for the slowdown of the growth rate. The model could fit the growth of lung metastasis visualized by magnetic resonance imaging. As a non-trivial outcome from this analysis, the model predicted that the net growth of two neighbouring tumour lesions that enter in contact is considerably impaired (of $31\% \pm 1.5\%$, mean \pm standard deviation), as compared to the growth of two independent tumours. Together, our results have implications for theories of metastatic development and suggest that global dynamics of metastasis development is dependent on spatial interactions between metastatic lesions.

Modeling spontaneous metastasis following surgery: an in vivo-in silico approach [6]

Team participant: Sebastien Benzekry

Other participants: Amanda Tracz, Michalis Matri, Ryan Corbelli, Dominique Barbolosis, John Ebos (Buffalo University)

Rapid improvements in the detection and tracking of early-stage tumor progression aim to guide decisions regarding cancer treatments as well as predict metastatic recurrence in patients following surgery. Mathematical models may have the potential to further assist in estimating metastatic risk, particularly when paired with in vivo tumor data that faithfully represent all stages of disease progression. Herein we describe mathematical analysis that uses data from mouse models of spontaneous metastasis developing after surgical removal of orthotopically implanted primary tumors. Both presurgical (primary tumor) and postsurgical (metastatic) growth was quantified using bioluminescence and was then used to generate a mathematical formalism based on general laws of the disease (i.e. dissemination and growth). The model was able to fit and predict pre-/post-surgical data at the level of the individual as well as the population. Our approach also enabled retrospective analysis of clinical data describing the probability of metastatic relapse as a function of primary tumor size. In these data-based models, inter-individual variability was quantified by a key parameter of intrinsic metastatic potential. Critically, our analysis identified a highly nonlinear relationship between primary tumor size and postsurgical survival, suggesting possible threshold limits for the utility of tumor size as a predictor of metastatic recurrence. These findings represent a novel use of clinically relevant models to assess the impact of surgery on metastatic potential and may guide optimal timing of treatments in neoadjuvant (presurgical) and adjuvant (postsurgical) settings to maximize patient benefit.

Migration and orientation of endothelial cells on micropatterned polymers: A simple model based on classical mechanics [11]**Team participants:** Thierry Colin, Clair Poignard, Olivier Saut**Other participants:** Julie Joie, Marie-Christine Durrieu (IMB - Institut de Mathématiques de Bordeaux), Yifeng Lei (French Institute of Health and Medical Research, Paris)

Understanding the endothelial cell migration on micropatterned polymers, as well as the cell orientation is a critical issue in tissue engineering, since it is the preliminary step towards cell polarization and that possibly leads to the blood vessel formation. In this work, we derive a simple agent-based model to describe the migration and the orientation of endothelial cells seeded on bioactive micropatterned polymers. The aim of the modeling is to provide a simple model that corroborates quantitatively the experiments, without considering the complex phenomena inherent to cell migration. Our model is obtained thanks to a classical mechanics approach based on experimental observations. Even though its simplicity, it provides numerical results that are quantitatively in accordance with the experimental data, and thus our approach can be seen as a preliminary way towards a simple modeling of cell migration.

7.3.2. Tumor-host crosstalk**Host age is a systemic regulator of gene expression impacting cancer progression [3]****Team participant:** Sebastien Benzekry**Other participants:** Afshin Beheshti, Lili Ma, Philip Hahnfeldt, Lynn Hlatky (CCSB - Center of Cancer and Systems Biology), J. Tyson McDonald (University of Houston), Michael Peluso (Cancer Risk Factor Branch, Molecular Biology Laboratory)

Aging is the major determinant of cancer incidence, which, in turn, is likely dictated in large part by processes that influence the progression of early subclinical (occult) cancers. However, there is little understanding of how aging informs changes in aggregate host signaling that favor cancer progression. In this study, we provide direct evidence that aging can serve as an organizing axis to define cancer progression-modulating processes. As a model system to explore this concept, we employed adolescent (68 days), young adult (143 days), middle-aged (551 days), and old (736 days) C57BL/6 mice as syngeneic hosts for engraftment of Lewis lung cancer to identify signaling and functional processes varying with host age. Older hosts exhibited dysregulated angiogenesis, metabolism, and apoptosis, all of which are associated with cancer progression. $TGF\beta 1$, a central player in these systemic processes, was downregulated consistently in older hosts. Our findings directly supported the conclusion of a strong host age dependence in determining the host tumor control dynamic. Furthermore, our results offer initial mechanism-based insights into how aging modulates tumor progression in ways that may be actionable for therapy or prevention.

Capturing the Driving Role of Tumor-Host Crosstalk in a Dynamical Model of Tumor Growth [4]**Team participant:** Sebastien Benzekry**Other participants:** Afshin Beheshti, Philip Hahnfeldt, Lynn Hlatky (CCSB - Center of Cancer and Systems Biology)

In 1999, Hahnfeldt et al. proposed a mathematical model for tumor growth as dictated by reciprocal communications between tumor and its associated vasculature, introducing the idea that a tumor is supported by a dynamic, rather than a static, carrying capacity. In this original work, the carrying capacity was equated with the variable tumor vascular support resulting from the net effect of tumor-derived angiogenesis stimulators and inhibitors. This dynamic carrying capacity model was further abstracted and developed in our recent publication to depict the more general situation where there is an interaction between the tumor and its supportive host tissue; in that case, as a function of host aging. This allowed us to predict a range of host changes that may be occurring with age that impact tumor dynamics. More generally, the basic formalism described here can be (and has been), extended to the therapeutic context using additional optimization criteria. The model depends on three parameters: one for the tumor cell proliferation kinetics, one for the stimulation of the stromal support, and one for its inhibition, as well as two initial conditions. We describe here the numerical method to estimate these parameters from longitudinal tumor volume measurements.

7.3.3. Metronomic oncology

Metronomic Reloaded: Theoretical Models Bringing Chemotherapy into the Era of Precision Medicine [5]

Team participant: Sebastien Benzekry

Other participants: Eddy Pasquier, Dominique Barbolosi, Joseph Ciccolini, Nicolas André (CRO2 - Centre de recherches en oncologie biologique et oncopharmacologie), Bruno Lacarelle (Clinical Pharmacokinetics), Fabrice Barlési (Service d'Oncologie Multidisciplinaire et d'Innovations Thérapeutiques)

Oncology has benefited from an increasingly growing number of groundbreaking innovations over the last decade. Targeted therapies, biotherapies, and the most recent immunotherapies all contribute to increase the number of therapeutic options for cancer patients. Consequently, substantial improvements in clinical outcomes for some disease with dismal prognosis such as lung carcinoma or melanoma have been achieved. Of note, the latest innovations in targeted therapies or biotherapies do not preclude the use of standard cytotoxic agents, mostly used in combination. Importantly, and despite the rise of bioguided (a.k.a. precision) medicine, the administration of chemotherapeutic agents still relies on the maximum tolerated drug (MTD) paradigm, a concept inherited from theories conceptualized nearly half a century ago. Alternative dosing schedules such as metronomic regimens, based upon the repeated and regular administration of low doses of chemotherapeutic drugs, have emerged as possible strategies to improve response rates while reducing toxicities. The recent changes in paradigm in the way we theorize cancer biology and evolution, metastatic spreading and tumor ecology, alongside the recent advances in the field of immunotherapy, have considerably strengthened the interest for metronomic approaches. This work aims at reviewing the recent evolutions in the field of theoretical biology of cancer and computational oncology, with a focus on the consequences these changes have on the way we administer chemotherapy. In particular, a step towards developing adaptive dosing should help to further optimize the efficacy of metronomic therapy. There is a rising trend to establish personalized medicine in oncology. Developing extensive bio-guided strategies for decision-making in the choice of drugs to be administered is now a common practice at the bedside. Similarly, developing extensive model-guided strategies for decision-making in refining dosing and scheduling should be undertaken to achieve precision medicine in oncology.

7.3.4. Protein-protein interaction networks

Design principles for cancer therapy guided by changes in complexity of protein-protein interaction networks [7]

Team participant: Sebastien Benzekry

Other participants: Jack A Tuszyński (Alberta University), Edward Rietman, Giannoula Lakka Klement (Newman-Lakka Institute)

The ever-increasing expanse of online bioinformatics data is enabling new ways to, not only explore the visualization of these data, but also to apply novel mathematical methods to extract meaningful information for clinically relevant analysis of pathways and treatment decisions. One of the methods used for computing topological characteristics of a space at different spatial resolutions is persistent homology. This concept can also be applied to network theory, and more specifically to protein-protein interaction networks, where the number of rings in an individual cancer network represents a measure of complexity. Results: We observed a linear correlation of $R = -0.55$ between persistent homology and 5-year survival of patients with a variety of cancers. This relationship was used to predict the proteins within a protein-protein interaction network with the most impact on cancer progression. By re-computing the persistent homology after computationally removing an individual node (protein) from the protein-protein interaction network, we were able to evaluate whether such an inhibition would lead to improvement in patient survival. The power of this approach lied in its ability to identify the effects of inhibition of multiple proteins and in the ability to expose whether the effect of a single inhibition may be amplified by inhibition of other proteins. More importantly, we illustrate specific examples of persistent homology calculations, which correctly predict the survival benefit observed effects in clinical trials using inhibitors of the identified molecular target. Conclusions: We propose that computational approaches such as persistent homology may be used in the future for selection of molecular therapies in clinic.

The technique uses a mathematical algorithm to evaluate the node (protein) whose inhibition has the highest potential to reduce network complexity. The greater the drop in persistent homology, the greater reduction in network complexity, and thus a larger potential for survival benefit. We hope that the use of advanced mathematics in medicine will provide timely information about the best drug combination for patients, and avoid the expense associated with an unsuccessful clinical trial, where drug(s) did not show a survival benefit.

7.4. Other new results

Superconvergent Cartesian Methods for Poisson type Equations in 2D-domains [21]

Team participants: Olivier Gallinato, Clair Poignard

In this work, we present three superconvergent Finite Difference methods on Cartesian grids for Poisson type equations with Dirichlet, Neumann or Robin conditions. Our methods are based on finite differences and high-order discretizations of the Laplace operator, to reach the superconvergence properties, in the sense that the first-order (and possibly the second-order) derivatives of the numerical solution are computed at the same order as the solution itself. We exhibit the numerical conditions that have to be fulfilled by the schemes to get such superconvergences and extensively illustrate our purpose by numerical simulations. We conclude by applying our method to a free boundary problem for cell protrusion formation recently proposed by the authors and colleagues. Note that quasistatic Stefan-like problem can be accurately solved by our methods.

Adaptive radiotherapy in routine: The radiation oncologist's point of view [14]

Team participant: Olivier Saut

Other participants: Bénédicte Henriques de Figueiredo, Adeline Petit, Paul Sargos, Guy Kantor, Claudia Pouypoudat, Christina Zacharitou, Mikael Antoine (Institut Bergonié, radiology department)

Adaptive radiotherapy is defined as all processes leading to the modification of a treatment plan on the basis of patient-specific variations observed during the course of a treatment. This concept is currently of particular relevance due to the development of onboard volumetric imaging systems, which allow for daily viewing of variations in both tumour and organs at risk in terms of position, shape or volume. However, its application in routine clinical practice is limited due to the demanding nature of the processes involved (re-delineation and replanning) and increased dependence on available human resources. Even if "online" strategies, based on deformable image registration (DIR) algorithms, could lead to a reduction in both work and calculation time, for the moment their use is limited to the research field due to uncertainties surrounding the validity of results gathered. Other strategies without DIR can be used as "offline" or "hybrid offline-online" strategies that seem to offer a compromise between time consumption and therapeutic gain for the patient.

MORPHEME Project-Team

6. New Results

6.1. Exact continuous penalties for ℓ_2 - ℓ_0 minimization

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Gilles Aubert.

We consider the following ℓ_0 -regularized least squares problem

$$\hat{x} \in \arg \min_{x \in \mathbb{R}^N} G_{\ell_0}(x) := \frac{1}{2} \|Ax - d\|^2 + \lambda \|x\|_0, \quad (11)$$

where $A \in \mathbb{R}^{M \times N}$, $d \in \mathbb{R}^M$ represents the data and $\lambda > 0$ is an hyperparameter characterizing the trade-off between data fidelity and sparsity. This problem finds a wide range of applications in signal/image processing, learning and coding areas among many others. We proposed a unified framework for exact continuous penalties approximating the ℓ_0 -norm. In other words, we are concerned by the design of a class of continuous relaxations of G_{ℓ_0} , preserving all its global minimizers, and for which any local minimal point is also one of the initial functional. Hence, we highlight five *necessary and sufficient* conditions on the continuous penalty approximating the ℓ_0 -norm ensuring that the minimizers of the underlying continuous relaxation of G_{ℓ_0} are consistent with those of G_{ℓ_0} . However, some local minimizer of the relaxed functional are not minimizer of G_{ℓ_0} which is an interesting point for such highly non-convex functional. This work offers a new way to compare penalties approximating the ℓ_0 -norm. Finally, it is worth noting that the CEL0 penalty [1], [14], [17] is the inferior limit of the obtained class of penalties and seems to be the best choice to do in order to obtained an equivalent continuous reformulation of (1).

6.2. Application of the Continuous Exact ℓ_0 relaxation to Channel and DOA sparse estimation problems

Participants: Emmanuel Soubies, Laure Blanc-Féraud.

This work is made in collaboration with Adilson Chinatto, Cynthia Junqueira, João M. T. Romano (University of Campinas, Brazil) and Pascal Larzabal, Jean-Pierre Barbot (ENS Cachan, SATIE Lab).

This work is devoted to two classical sparse problems in array processing: Channel estimation and DOA (Direction Of Arrivals) estimation. We show how our results on ℓ_0 optimization [1], [14], [17] can be used, at the same computational cost, in order to obtain improvement in comparison with ℓ_1 optimization (usually used) for sparse estimation. Moreover, for the DOA case, we show that our analysis conducted in the Single Measurement Vector (SMV) case [1] can be generalized to the Multiple Measurement Vectors (MMV) case. In that case, the variable x is not a vector of \mathbb{R}^N but a matrix of $\mathbb{R}^{N \times K}$ where N is the signal length and K the number of snapshots. Hence, one wants to apply sparsity to the rows of x , i.e. x must have a small number of nonzero rows, instead of applying the sparsity on all the components of x . This results in a row-structured sparsity penalty which is modelled using a mixed ℓ_2 - ℓ_0 norm.

Finally, numerical experiments demonstrate the efficiency of the proposed approach compared to classical methods as ℓ_1 relaxation, Iterative Hard Thresholding or MUSIC algorithms and that it can reach the Cramer Rao Bound in some cases [4].

6.3. From TIRF microscope calibration to 3D biological reconstructions

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Sébastien Schaub, Gilles Aubert.

This work is made in collaboration with Agata Radwanska, Ellen Van Obberghen-Schilling (iBV).

Total Internal Reflection Fluorescence microscopy (TIRF) is a method of choice to visualize membrane-substrate interactions. The principle of this device relies on the total internal reflection phenomenon generating an evanescent wave capable of producing a selective excitation of the dye molecules within a single layer of 100 to 500nm. The fast decay of the evanescent wave varies with respect to the incident angle of the light beam. Hence, intensity variations on TIRF images, occurring when changing the incident angle, are, in part, due to the axial positions of the observed structures. While a direct interpretation of Multi-Angle TIRF (MA-TIRF) images in terms of axial structure positions is not an easy task, reconstruction algorithms can be dedicated to compute a quantitative depth map with high axial resolution. However, the success of such reconstruction methods strongly depends on the system calibration.

We have proposed a pipeline for MA-TIRF calibration. Considering back focal plane (BFP) images of several solutions differing by their refractive indices, we validate the theoretical relation linking the tension applied to the galvanometric mirror (which controls the laser beam orientation) and the incident angle of the beam on the specimen. Then it is crucial to verify if the simple exponential decaying model of the evanescent wave is sufficient to describe our setup. To this end we propose to build a phantom sample (for which the geometry is known) using a large lens placed into a homogeneous fluorescent solution (Fig. 1 top). Based on a least square estimation, we showed a good agreement between the estimated slope of the lens (we assume the lens to be linear near the border) and the expected one up to 400nm depth (Fig. 1 bottom-left). To complete the validation procedure, we use a sample for which the structures of interest are labeled using two different fluorescent proteins sensitive to different wavelengths and emitting respectively green and red fluorescence. Then, using standard variational approaches, we obtain a co-localization of the reconstructed structures with a precision around 30-40nm (Fig. 1 bottom-middle) over at least 170nm depth showing the precision of the method. Finally, once this calibration step is achieved, we perform color-coded depth representation of 3D biological structures living in the vicinity of the cell membrane (Fig. 1 bottom-right). All these experiments have been made on an experimental TIRF system developed at iBV lab in Valrose.

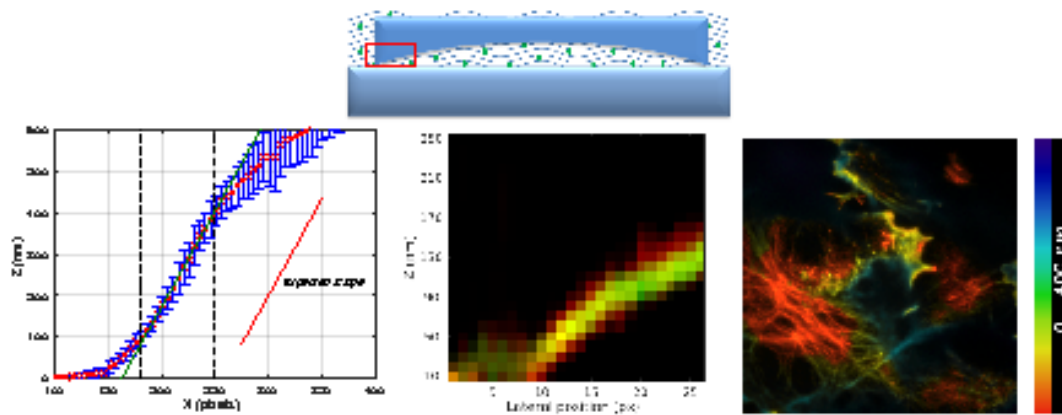


Figure 1. Top: Phantom sample constructed from a large lens and an homogeneous fluorescent solution. The red rectangle represents the observed region through the MA-TIRF setup. Bottom: estimated (green) and expected (red) slope of the lens within the red zone of the top figure (left), results of the co-localization experiment along a XY-line (middle) and a color-coded depth representation of a 3D biological reconstruction (right).

6.4. Phase estimation in Differential Interference Contrast (DIC) microscopy

Participants: Lola-Xiomara Bautista Rozo, Laure Blanc-Féraud.

We present a gradient-based optimization method for the estimation of a specimen phase function from polychromatic DIC images. The method minimizes the sum of a nonlinear least-squares discrepancy measure and a smooth approximation of the total variation. A new formulation of the gradient and a recent updating rule for the choice of the step size are both exploited to reduce computational time. Numerical simulations on two computer-generated objects show significant improvements, both in efficiency and accuracy, with respect to a more standard choice of the step size.

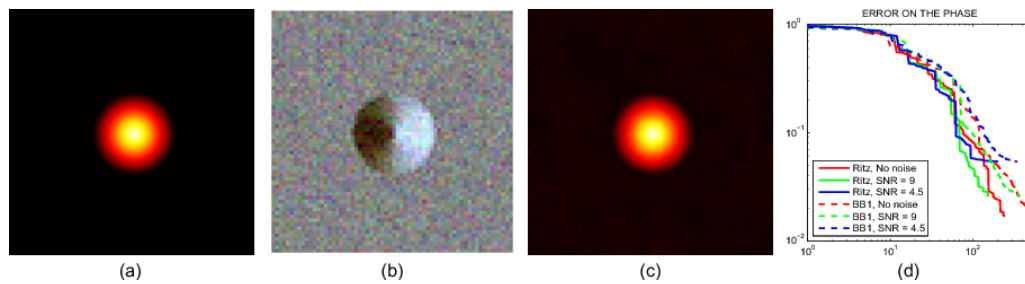


Figure 2. Data and results for the cone object. From left to right: true object, noisy DIC color image taken at angle $\tau_0 = 0^\circ$ and $SNR = 4.5$, reconstructed phase and the relative error versus the number of iterations.

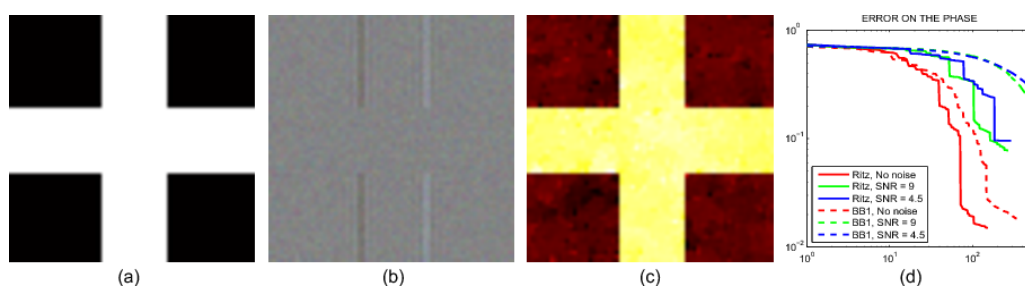


Figure 3. Data and results for the cross object. From left to right: true object, noisy DIC color image taken at angle $\tau_0 = 0^\circ$ and $SNR = 4.5$, reconstructed phase with and the relative error versus the number of iterations.

6.5. Spatio-temporal registration of 3D microscopy image sequences of Arabidopsis floral meristems

Participants: Gaël Michelin, Grégoire Malandain.

This work is made in collaboration with Léo Guignard and Christophe Godin (Virtual Plants), within the Morphogenetics Inria Project Lab.

The shoot apical meristem (SAM) is at the origin of all the plant above-ground organs (including stems, leaves and flowers) and is a biological object of interest for the understanding of plant morphogenesis. The quantification of tissue growth at a cellular level requires the analysis of 3D microscopic image sequences of developing meristems. To address inter-individual variability, it is also required to compare individuals. This obviously implies the ability to process inter-individual registration, i.e. to compute spatial and temporal correspondences between sequences from different meristems.

In [8], we propose a spatial registration method dedicated to microscopy floral meristem (FM) images, based on the registration of both the outer and the inner surfaces of the L1 layer (the epidermal cell layer). A given meristem (one timepoint) can be compared to a sequence (several timepoints) of another meristem: the goodness-of-fit criterion allows to identify the best corresponding time-point in this sequence of another individual, achieving the temporal registration (see figure 4). Since the morphological deformations remain important between successive images of a sequence, images interpolation between time-points is also performed in order to refine the sequence temporal resolution and thus to ensure a precise temporal registration.

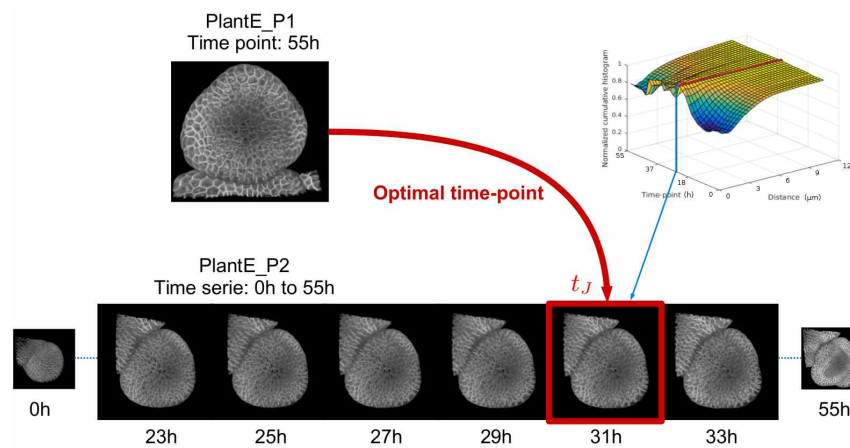


Figure 4. Inter-individual temporal registration result with 3D views of the registered meristem and the interpolated movie at several time-points.

6.6. Epidermal cell layer thickness variability in Arabidopsis floral meristems

Participants: Gaël Michelin, Grégoire Malandain.

This work is made in collaboration with Yassin Refahi (Sainsbury Lab., University of Cambridge) and Jan Traas (ENS Lyon), within the Morphogenetics Inria Project Lab.

Flowers from the same species display a great robustness in their global shape and their developing stage can be theoretically identifiable to their size. The cells in epidermal (L1) and sub-epidermal (L2) layers of the floral meristem divide anticlinally, i.e. in a sideways fashion that ensures that L1 and L2 remain distinct. Thus a goodness-of-fit criterion on L1 and L2 layers is considered as an adequate registration quality measure in the inter-individual spatio-temporal registration framework developed in [8].

The aim of the present work is to investigate the variability of L1 layer thickness over development stages of an individual and between individuals. The study results may impact the way we process the inter-individual spatial registration. Therefore we measured the thickness distribution (histogram) of the L1 cells and we plotted the distribution of cells thickness (see figure 5) on images provided from three distinct floral meristems at acquisition time-points. Our results tend towards showing that L1 thickness increases over time non-uniformly, with a higher L1 thickness on sepals for advanced developing stages. We also observed an inter-individual thickness variability of about 15% for developing floral meristems at close developing stages. Future investigations will consist in taking a larger set of data to assess our first observations, in providing a biological interpretation of these observations and in using this knowledge to propose a refined spatial registration method.

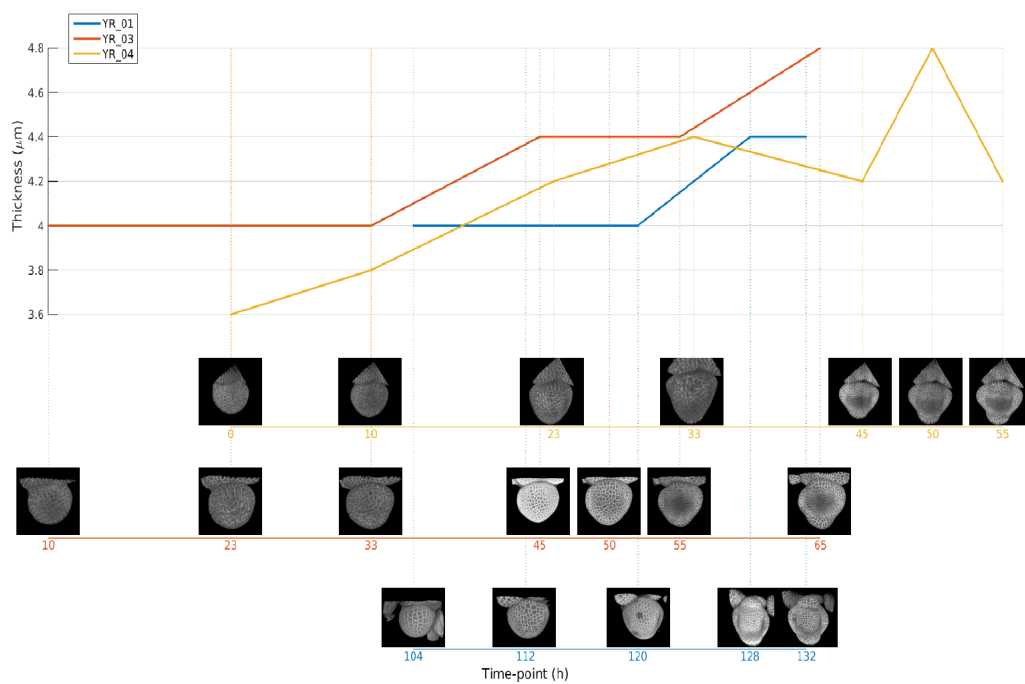


Figure 5. Epidermal cell layer thickness distribution over floral meristem development. Each row relates the measures at different developing times of a floral meristem.

6.7. Statistical Characterization, Modelling and Classification of Morphological Changes in *imp* Mutant *Drosophila* Gamma Neurons

Participants: Agustina Razetti, Caroline Medioni, Florence Besse, Xavier Descombes.

In *Drosophila* brain, gamma neurons in the mushroom body are involved in higher functions such as olfactory learning and memory. During metamorphosis, they undergo remodelling after which they adopt their adult shape. Some mutations alter remodelling and therefore neuronal final morphology, causing behavioural dysfunctions. The RNA binding protein *Imp*, for example, was shown to control this remodelling process at least partly by regulating profilin expression. This work aims at precisely characterizing the morphological changes observed upon *imp* knockdown in order to further understand the role of this protein. We developed a methodological framework that consists in the selection of relevant morphological features (axon length and shape and branch length distribution and density), their modelling and parameter estimation. We thus perform a statistical comparison and a likelihood analysis to quantify similarities and differences between wild type and mutated neurons. The data was a set of 3D images showing a single neuron taken with a confocal microscope and provided by F. Besse group, IBV. The workflow from raw data to the likelihood analysis is summarized on figure 6. We show that *imp* mutant neurons can be classified into two phenotypic groups (called *Imp L* and *Imp Sh*) that differ in several morphological aspects. We also demonstrate that, although *Imp L* and wild-type neurons show similarities, branch length distribution is discriminant between these populations. Finally, we study biological samples in which Profilin was reintroduced in *imp* mutant neurons, and show that defects in main axon and branch lengths are partially suppressed.

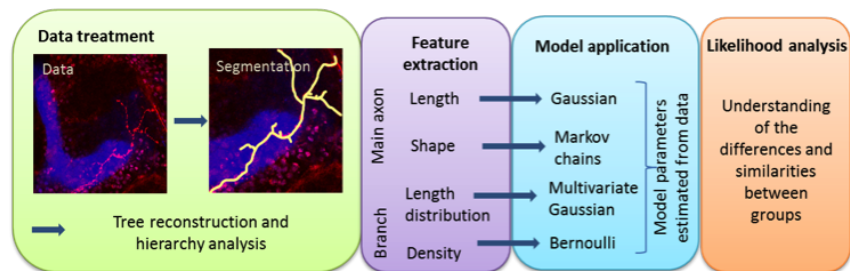


Figure 6. Summary of the workflow from raw data to the likelihood analysis.

6.8. Genome-wide search for factors that control the assembly of RNA granules

Participants: Wei Shen, Nicolas Cedilnik, Florence Besse, Xavier Descombes.

This work has been done in collaboration with Fabienne De Graeve from iBV

In vivo, mRNAs are packaged together with regulatory proteins into ribonucleoprotein particles (RNP) that control their fate and undergo extensive remodeling in response to developmental cues or environmental stresses. Cytoplasmic RNPs of different sizes, composition and regulatory properties have been described, including large macromolecular complexes such as P-bodies, stress granules, or germ cell granules. We aim at studying the different RNA granules distribution within the cytoplasm depending on genomic factors.

Before considering a spatial statistics analysis of the granules, it is necessary to detect them on confocal microscopy images of the cells. Therefore, we have studied a first pipeline for detecting these granules in confocal microscopy images. We have marked cells with DAPI for detecting nuclei. These nuclei are then classified into "dead" or "alive" by a support vector machine (SVM) using intensity and shape criteria. In the

second step we consider GFP marked images to segment the cytoplasm and detect the granules within the cytoplasm. The cytoplasm segmentation is performed using an active contour whereas the granule detection is based on a marked point process model optimized by the multiple births and cut algorithm.

The full pipeline has been validated on a few samples from a pilot study. The next step will consists of a validation on the full study before considering a genome-wide screening.

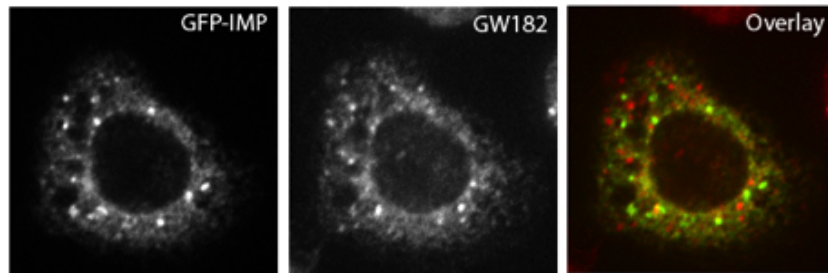


Figure 7. GFP-IMP particles are distinct from P-bodies. S2R+ cells expressing GFP-IMP fusions (left: green in the overlay) and stained with α -GW182 antibodies (middle: red in the overlay). GW182 is a well-described marker of P-bodies. Experiment performed in F. Besse lab at iBV (unpublished).

6.9. Cells detection using segmentation competition

Participants: Sen Wang, Emmanuel Soubies, Xavier Descombes.

Marked point processes have proved to be very efficient for segmenting a collection of objects from digital images. The multiple birth and death algorithm provides an optimization framework that allows reasonable time computation. This algorithm has to be embedded in a simulated annealing framework which involves parameters tuning (initial temperature and cooling scheme). This tedious task can be overcome considering a graph cut algorithm instead of the probabilistic death step. The algorithm then consists in successively adding new random objects in the configuration and selecting the most relevant using the graph cut algorithm. In the graph construction a node is associated to each object. In the original algorithm proposed by [21] the regularity condition imposed by the graph cut prevents to consider attractive interactions such as clustering or alignment constraints, which restricts the model to repulsive properties such as non overlap between objects. To overcome this restriction we have investigated new graph constructions by considering nodes defined by clusters of interacting objects. Different strategies have been compared to avoid being trapped in local minima defined by clusters while minimizing the number of required iterations. We have applied this new algorithm on different bioimaging problems such as axon extraction or cells detection (see figure 8).

6.10. Vesicles trajectory detection and analysis

Participant: Xavier Descombes.

This work has been done in collaboration with Maximilian Furthauer and Thomas Juan from iBV.

In many species, the left right asymmetry of organs location is initiated in a ciliated cavity called Kupffer's vesicle in zebrafish. The cills beating induce a non symmetrical flow in the cavity that can be studied by following the trajectory of exovesicles in the Kupffer's vesicle. The goal of this project is to automatically track these exovesicles and to perform a statistical analysis of these trajectories in different conditions.



Figure 8. Axon detection on a synthetic image without (left) and with (middle left) attractive interactions and on real image (middle right) with attractive interactions (right).

We consider 2D time sequences of images. To extract the vesicles from the time sequence we first remove the background by subtracting a local time mean. We then detect the cell border using an active contour computed on the spatial derivative of the images. The vesicles are then simply detected using a threshold followed by a morphological opening to remove the noise. The trajectory are finally obtained using a morphological closing in time. We aim at statistically comparing populations. In order to aggregate trajectories from several samples, we project the datasets into the same space using a continuous transformation of each cell into a reference disk. We thus project all the obtained trajectories from a given population into this disk. We compute the speed vector on each time point of each detected trajectory. To obtain a dense representation of the norm and the orientation of the vector speed in the reference disk, we extrapolate the obtained vector speed to a regular lattice with a Gaussian Markov random field. Finally, we obtain two spatial maps of respectively the norm and the orientation of the speed [9](#).

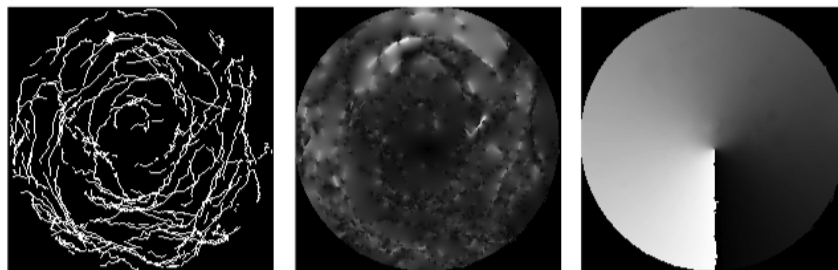


Figure 9. Exovesicles trajectories detected on the control population (left), the speed norm map (middle) and the angle map (right) showing an anticlockwise movement (black=0, white= 2π).

6.11. Extraction and Analysis of the Vascular Network to Classify and Grade Kidney Tumors in Histological Imaging

Participants: Alexis Zubiolo, Eric Debreuve, Xavier Descombes.

This work is made in collaboration with Philippe Pognonec (Team TIRO, CEA/UNS), Damien Ambrosetti (Histopathology department, CHU Pasteur, Nice).

The renal carcinoma is the most frequent type of kidney cancer (between 90% and 95% of all cases). Twelve classes of carcinoma can be distinguished, among which the clear cell carcinoma (CCRCC) and the papillary carcinoma (PRCC) are the two most common (75% and 10% of the cases, respectively). After the carcinoma has been diagnosed, the tumor is ablated and prepared for histological examination (fixation, staining, slicing, observation with a microscope). Along with genetic tests and protein reactions, the histological study allows to classify and grade the tumor in order to make a prognosis and to take decisions for the subsequent patient treatment. Digital histology is a recent domain (routinely, histological slices are studied by MDs directly on the microscope). The pioneer works deal with the automatic analysis of cells. However, one crucial factor for carcinoma classification is the structure of the vascular network. Coarsely, CCRCC is characterized by a “fishnet” structure while the PRCC has a tree-like structure.

In this context, our goal was to extract the vascular network from a given histological slice, compute features of the underlying graph structure, and classify the tumor into CCRCC or PRCC based on these features. The histological images being huge (typically, 100k x 100k pixels), they must be split into tiles (with some overlap to ease the combination of results) and processed tile-wise. The first step is to combine the color channels so that the vessels are as highlighted as possible. Then, the vascular network is detected by a processing pipeline including tailored, Gabor-like filtering, thresholding, and extraction of the skeleton. Small gaps in the skeleton are filled and some pruning is performed. Finally, the skeleton is converted to a graph representation. Based on the medical interpretation procedure, we focused our analysis of the graph on the following elements: the number of terminal and junction nodes, and the terminal branches. We proposed to compute the ratio between the number of terminal nodes and the number of junctions (T/J ratio), and the average length of terminal branches. Both features seem to be adapted to classification, especially the T/J ratio which, on the fairly small database of cases we currently have, exhibits an average value 65% higher for PRCC.

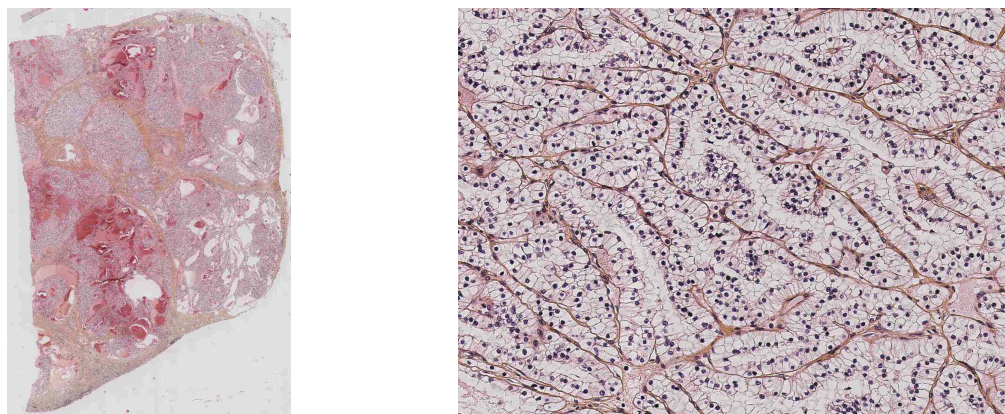


Figure 10. A histological slice through a kidney tumor: the whole slice (left) and a close-up (right).

MYCENAE Project-Team

6. New Results

6.1. Numerical and theoretical studies of slow-fast systems with complex oscillations

6.1.1. *Canard-Mediated (De)Synchronization in Coupled Phantom Bursters*

Participants: Elif Köksal Ersöz, Mathieu Desroches, Maciej Krupa, Frédérique Clément.

In [32], we study canard-mediated transitions in mutually coupled phantom bursters. We extend a multiple-timescale model which provides a sequence of dynamic events, i.e. transition from a frequency modulated relaxation cycle to a quasi-steady state and resumption of the relaxation regime through small amplitude oscillations. Folded singularities and associated canard solutions have a particular impact on the dynamics of the original system, which consists of two feedforward coupled FitzHugh-Nagumo oscillators, where the slow subsystem (regulator) controls the periodic behavior of the fast subsystem (secretor). We first investigate the variability in the dynamics depending on the canard mechanism that occurs near the folded singularities of the 4D secretor-regulator configuration. Then, we introduce a second secretor and focus on the slow-fast transitions in the presence of a linear coupling between the secretors. In particular, we explore the impact of the relationship between the canard structures and the coupling on patterns of synchronization and desynchronization of the collective dynamics of the resulting 6D system. We identify two different sources of desynchronization induced by canards, near a folded-saddle singularity and a folded-node singularity, respectively.

Part of these results have also been presented as posters at the *SIAM Conference on Applications of Dynamical Systems* (Snowbird, May 17-21, 2015) and *1st International Conference on Mathematical Neuroscience* (Antibes Juan les Pins, June 8-10-2015).

6.1.2. *Mixed-Mode Oscillations in a piecewise linear system with multiple time scale coupling*

Participants: Soledad Fernández García, Maciej Krupa, Frédérique Clément.

We analyze a four dimensional slow-fast piecewise linear system with three time scales presenting Mixed-Mode Oscillations. The system possesses an attractive limit cycle along which oscillations of three different amplitudes and frequencies can appear, namely, small oscillations, pulses (medium amplitude) and one surge (largest amplitude). In addition to proving the existence and attractiveness of the limit cycle, we focus our attention on the canard phenomena underlying the changes in the number of small oscillations and pulses. We analyze locally the existence of secondary canards leading to the addition or subtraction of one small oscillation and describe how this change is globally compensated for or not with the addition or subtraction of one pulse.

6.1.3. *Noise-induced canard and mixed-mode oscillations in large stochastic networks with multiple timescales*

Participants: Jonathan Touboul, Maciej Krupa, Mathieu Desroches.

We investigate in [28] the dynamics of large stochastic networks with different timescales and nonlinear mean-field interactions. After deriving the limit equations for a general class of network models, we apply our results to the celebrated Wilson-Cowan system with two populations with or without slow adaptation, paradigmatic example of nonlinear mean-field network. This system has the property that the dynamics of the mean of the solution exactly satisfies an ODE. This reduction allows to show that in the mean-field limit and in multiple populations with multiple timescales, noise induces canard explosions and Mixed-Mode Oscillations on the mean of the solution. This sheds new light on the qualitative effects of noise and sensitivity to precise noise values in large stochastic networks. We further investigate finite-sized networks and show that systematic differences with the mean-field limits arise in bistable regimes (where random switches between different attractors occur) or in mixed-mode oscillations, where the finite-size effects induce early jumps due to the sensitivity of the attractor.

6.1.4. *Canard explosion in delayed equations with multiple timescales, applications to the delayed Fitzhugh-Nagumo system*

Participants: Maciej Krupa, Jonathan Touboul.

In two contributions, we investigated theoretically the presence of canard explosions of delayed differential equations, and have applied these results to the FitzHugh-Nagumo neuronal model.

- In [21] we analyze canard explosions in delayed differential equations with a one-dimensional slow manifold. This study is applied to explore the dynamics of the van der Pol slow-fast system with delayed self-coupling. In the absence of delays, this system provides a canonical example of a canard explosion. We show that as the delay is increased a family of ‘classical’ canard explosions ends as a Bogdanov-Takens bifurcation occurs at the folds points of the S-shaped critical manifold.
- Motivated by the dynamics of neuronal responses, we analyze in [21] the dynamics of the Fitzhugh-Nagumo slow-fast system with delayed self-coupling. Beyond the regime of small delays, delays significantly enrich the dynamics, leading to mixed-mode oscillations, bursting and chaos. These behaviors emerge from a delay-induced subcritical Bogdanov-Takens instability arising at the fold points of the S-shaped critical manifold. Underlying the transition from canard-induced to delay-induced dynamics is an abrupt switch in the nature of the Hopf bifurcation.

6.1.5. *Canard-induced loss of stability across a homoclinic bifurcation*

Participants: Mathieu Desroches, Jean-Pierre Franoise, Lucile Megret.

In [16], we investigate the possibility of bifurcations which display a dramatic change in the phase portrait in a very small (on the order of 10^{-7} in the example presented here) change of a parameter. We provide evidence of existence of such a very rapid loss of stability on a specific example of a singular perturbation setting. This example is strongly inspired of the explosion of canard cycles first discovered and studied by E. Benoît, J.-L. Callot, F. Diener and M. Diener. After some presentation of the integrable case to be perturbed, we present the numerical evidences for this rapid loss of stability using numerical continuation. We discuss then the possibility to estimate accurately the value of the parameter for which this bifurcation occurs.

6.1.6. *Analysis of Interspike-Intervals for the General Class of Integrate-and-Fire Models with Periodic Drive*

Participant: Justyna Signerska-Rynkowska.

In [27], we study one-dimensional integrate-and-fire models of the general type $\dot{x} = F(t, x)$ and analyze properties of the firing map which iterations recover consecutive spike timings. We impose very weak constraints for the regularity of the function $F(t, x)$ e.g. often it suffices to assume that F is continuous. If additionally F is periodic in t , using mathematical study of the displacement sequence of an orientation preserving circle homeomorphism, we provide a detailed description of the regularity properties of the sequence of interspike-intervals and behaviour of the interspike-interval distribution.

6.1.7. *A geometric mechanism for mixed-mode bursting oscillations in a hybrid neuron model*

Participants: Justyna Signerska-Rynkowska, Jonathan Touboul, Alexandre Vidal.

In [35], we exhibit and investigate a new type of mechanism for generating complex oscillations featuring an alternation of small oscillations with spikes (MMOs) or bursts (MMBOs) in a class of hybrid dynamical systems modeling neuronal activity. These dynamical systems, called nonlinear adaptive integrate-and-fire neurons, combine nonlinear dynamics modeling input integration in a nerve cell with discrete resets modeling the emission of an action potential and the subsequent return to reversal potential. We show that presence of complex oscillations in these models relies on a fundamentally hybrid structure of the flow: invariant manifolds of the continuous dynamics govern small oscillations, while discrete resets govern the emission of spikes or bursts. The decomposition into these two mechanisms leads us to propose a purely geometrical interpretation of these complex trajectories, and this relative simplicity allows to finely characterize the MMO patterns through the study of iterates of the adaptation map associated with the hybrid system. This map is however

singular: it is discontinuous and has unbounded left- and right-derivatives. We apply and develop rotation theory of circle maps for this class of adaptation maps to precisely characterize the trajectories with respect to the parameters of the system. In contrast to more classical frameworks in which MM(B)Os were evidenced, the present geometric mechanism neither requires no more than two dimensions, does not necessitate to have separation of timescales nor complex return mechanisms.

Part of these results have also been presented as posters at the *SIAM Conference on Applications of Dynamical Systems* (Snowbird, May 17-21, 2015) and *1st International Conference on Mathematical Neuroscience* (Antibes Juan les Pins, June 8-10-2015).

6.2. Non conservative transport equations for cell population dynamics

6.2.1. Cell-kinetics based calibration of a multiscale model: application to cell population dynamics in ovarian follicles

Participants: Benjamin Aymard [ICL], Frédérique Clément, Danielle Monniaux [INRA], Marie Postel.

In [30], we present a strategy for tuning the parameters of a multiscale model of structured cell populations in which physiological mechanisms are embedded into the cell scale. This strategy allows one to cope with the technical difficulties raised by such models, that arise from their anchorage in cell biology concepts: localized mitosis, progression within and out of the cell cycle driven by time- and possibly unknown-dependent, and nonsmooth velocity coefficients. We compute different mesoscopic and macroscopic quantities from the microscopic unknowns (cell densities) and relate them to experimental cell kinetic indexes. We study the expression of reaching times corresponding to characteristic cellular transitions in a particle-like reduction of the original model. We make use of this framework to obtain an appropriate initial guess for the parameters and then perform a sequence of optimization steps subject to quantitative specifications. We finally illustrate realistic simulations of the cell populations in cohorts of interacting ovarian follicles.

6.2.2. Dimensional reduction of a multiscale cell population model

Participants: Frédérique Clément, Frédéric Coquel [CMAP], Marie Postel, Kim Long Tran.

We have designed a dimensional reduction of a multiscale structured cell population model, consisting of a system of 2D transport equations, into a system of twice as many 1D transport equations. The reduced model is obtained by computing the moments of the 2D model with respect to one space variable. The 1D solution is defined from the solution of the 2D model starting from an initial condition that is a Dirac mass in the direction removed by reduction. Long time properties of the 1D model solution are obtained in connection with properties of the support of the 2D solution for general case initial conditions. Finite volume numerical approximations of the 1D reduced model can be used to compute the moments of the 2D solution with satisfying accuracy. The numerical robustness is studied in the scalar case and a full scale vector case is presented.

6.3. Macroscopic limits of stochastic neural networks and neural fields

6.3.1. Pinwheel-Dipole configuration in cat visual cortex

Participants: Jérôme Ribot [CIRB], Alberto Romagnoni [CIRB], Chantal Milleret [CIRB], Daniel Bennequin [CIRB], Jonathan Touboul.

One fascinating aspect of the brain is its ability to process information in a fast and reliable manner. The functional architecture is thought to play a central role in this task, by encoding efficiently complex stimuli and facilitating higher level processing. In the early visual cortex of higher mammals, information is processed within functional maps whose layout is thought to underlie visual perception. The possible principles underlying the topology of the different maps, as well as the role of a specific functional architecture on information processing, is however poorly understood.

- In [25], we show that spatial frequency representation in cat areas 17 and 18 exhibits singularities around which the map organizes like an electric dipole potential. These singularities are precisely co-located with singularities of the orientation map: the pinwheel centers. We first show, using high resolution optical imaging, that a large majority (around 80%) of pinwheel centers exhibit in their neighborhood semi-global extrema in the spatial frequency map. These extrema created a sharp gradient that was confirmed with electrophysiological recordings. Based on an analogy with electromagnetism, a mathematical model of a dipolar structure is proposed, that was accurately fitted to optical imaging data for two third of pinwheel centers with semi-global extrema.
- Mathematically, this pinwheel-dipole architecture is fascinating. We demonstrated mathematically in [26] that two natural principles, local exhaustivity of representation and parsimony, would indeed constrain the orientation and spatial frequency maps to display co-located singularities around which the orientation is organized as a pinwheel and spatial frequency as a dipole. Moreover, using a computational model, we showed that this architecture allows a trade-off in the local perception of orientation and spatial frequency, but this would occur for sharper selectivity than the tuning width reported in the literature. We therefore re-examined physiological data and show that indeed the spatial frequency selectivity substantially sharpens near maps singularities, bringing to the prediction that the system tends to optimize balanced detection between different attributes.

These results shed new light on the principles at play in the emergence of functional architecture of cortical maps, as well as their potential role in processing information.

6.3.2. *Absorption properties of stochastic equations with Hölder diffusion coefficients*

Participants: Jonathan Touboul, Gilles Wainrib [ENS].

In [29], we address the absorption properties of a class of stochastic differential equations around singular points where both the drift and diffusion functions vanish. According to the Hölder coefficient alpha of the diffusion function around the singular point, we identify different regimes. Stability of the absorbing state, large deviations for the absorption time, existence of stationary or quasi-stationary distributions are discussed. In particular, we show that quasi-stationary distributions only exist for $\alpha < 3/4$, and for alpha in the interval $(3/4, 1)$, no quasi-stationary distribution is found and numerical simulations tend to show that the process conditioned on not being absorbed initiates an almost sure exponential convergence towards the absorbing state (as is demonstrated to be true for $\alpha = 1$). Applications of these results to stochastic bifurcations are discussed.

6.3.3. *On a kinetic FitzHugh-Nagumo model of neuronal network*

Participants: Stéphane Mischler [CEREMADE], Cristóbal Quiñinao [CIRB], Jonathan Touboul.

We investigate in [33] the existence and uniqueness of solutions of a McKean-Vlasov evolution PDE representing the macroscopic behavior of interacting Fitzhugh-Nagumo neurons. This equation is hypoelliptic, nonlocal and has unbounded coefficients. We proved existence of a solution to the evolution equation and non trivial stationary solutions. Moreover, we demonstrated uniqueness of the stationary solution in the weakly nonlinear regime. Eventually, using a semigroup factorisation method, we showed exponential nonlinear stability in the small connectivity regime.

6.4. Modeling of neurogenesis and brain development

6.4.1. *Lhx2 regulates the timing of β -catenin-dependent cortical neurogenesis*

Participants: Lea-Chia-Ling Hsu [Taipei], Sean Nama [Taipei], Yi Cui, Ching-Pu Chang [Taipei], Chia-Fang Wang [Taipei], Hung-Chih Kuo [Taipei], Jonathan Touboul, Shen-Ju Chou [Taipei].

The timing of cortical neurogenesis has a major effect on the size and organization of the mature cortex. The deletion of the LIM-homeodomain transcription factor *Lhx2* in cortical progenitors by Nestin-cre leads to a dramatically smaller cortex. In [19] we report that *Lhx2* regulates the cortex size by maintaining the cortical progenitor proliferation and delaying the initiation of neurogenesis. The loss of *Lhx2* in cortical progenitors results in precocious radial glia differentiation and a temporal shift of cortical neurogenesis. We further investigated the underlying mechanisms at play and demonstrated that in the absence of *Lhx2*, the Wnt/ β -catenin pathway failed to maintain progenitor proliferation. We developed and applied a mathematical model that reveals how precocious neurogenesis affected cortical surface and thickness. Thus, we concluded that *Lhx2* is required for β -catenin function in maintaining cortical progenitor proliferation and controls the timing of cortical neurogenesis.

6.4.2. Competition and boundary formation in heterogeneous media: Application to neuronal differentiation

Participants: Cristóbal Quiñinao [CIRB], Benoît Perthame [LJLL], Jonathan Touboul.

We analyze in [22] an inhomogeneous system of coupled reaction-diffusion equations representing the dynamics of gene expression during differentiation of nerve cells. The outcome of this developmental phase is the formation of distinct functional areas separated by sharp and smooth boundaries. It proceeds through the competition between the expression of two genes whose expression is driven by monotonic gradients of chemicals, and the products of gene expression undergo local diffusion and drive gene expression in neighboring cells. The problem therefore falls in a more general setting of species in competition within a non-homogeneous medium. We show that in the limit of arbitrarily small diffusion, there exists a unique monotonic stationary solution, which splits the neural tissue into two winner-take-all parts at a precise boundary point: on both sides of the boundary, different neuronal types are present. In order to further characterize the location of this boundary, we use a blow-up of the system and define a traveling wave problem parametrized by the position within the monotonic gradient: the precise boundary location is given by the unique point in space at which the speed of the wave vanishes.

6.4.3. Local homeoprotein diffusion can stabilize boundaries generated by graded positional cues

Participants: Cristóbal Quiñinao [CIRB], Alain Prochiantz [CIRB], Jonathan Touboul.

Boundary formation in the developing neuroepithelium decides on the position and size of compartments in the adult nervous system. In [23], we started from the French Flag model proposed by Lewis Wolpert, in which boundaries are formed through the combination of morphogen diffusion and of thresholds in cell responses. In contemporary terms, a response is characterized by the expression of cell-autonomous transcription factors, very often of the homeoprotein family. Theoretical studies suggest that this sole mechanism results in the formation of boundaries of imprecise shapes and positions. Alan Turing, on the other hand, proposed a model whereby two morphogens that exhibit self-activation and reciprocal inhibition, and are uniformly distributed and diffuse at different rates lead to the formation of territories of unpredictable shapes and positions but with sharp boundaries (the 'leopard spots'). Here, we have combined the two models and compared the stability of boundaries when the hypothesis of local homeoprotein intercellular diffusion is, or is not, introduced in the equations. We find that the addition of homeoprotein local diffusion leads to a dramatic stabilization of the positioning of the boundary, even when other parameters are significantly modified. This novel Turing/Wolpert combined model has thus important theoretical consequences for our understanding of the role of the intercellular diffusion of homeoproteins in the developmental robustness of and the changes that take place in the course of evolution.

6.4.4. Designing a mathematical model of the dynamics of progenitor cell populations in the mouse cerebral cortex

Participants: Marie Postel, Alice Karam [UPMC], Mérina Latbi [UPMC], Guillaume Pezeron [UPMC], Kim Long Tran, Frédérique Clément, Sylvie Schneider-Maunoury [UPMC].

The mammalian cortex is a laminar structure in the dorsal telencephalon, composed of distinct cell types with different spatial and temporal origins. Cortical projection neurons display different patterns of layering and connectivity that depend on their birth date. We have designed a multi-scale mathematical model of structured cell populations, taking into account three main cell types: apical progenitors (APs), intermediate progenitors (IPs) and neurons (N). APs self-renew and produce IPs that divide to give Ns. The main originality of this spatio-temporal model is to explicitly represent the different phases of the cell cycle, G1, S, G2 and M. Biological data from the experiments and from the literature provide values for parameters of the model (e.g. duration of each cell cycle phase and division rates for each cell type). The outputs of the model are interpretable in terms of cell kinetics (e.g. mitotic index, labelling index, cell numbers). They are adjusted to experimental observations by numerical simulation.

NEUROMATHCOMP Project-Team

6. New Results

6.1. Neural Networks as dynamical systems

6.1.1. *Periodic forcing of stabilized E-I networks: Nonlinear resonance curves and dynamics*

Participants: Romain Veltz, Terry Sejnowski [Salk Institute].

Inhibition stabilized networks (ISNs) are neural architectures with strong positive feedback among pyramidal neurons balanced by strong negative feedback from inhibitory interneurons, a circuit element found in the hippocampus and the primary visual cortex. In their working regime, ISNs produce damped oscillations in the γ -range in response to inputs to the inhibitory population. In order to understand the properties of interconnected ISNs, we investigated periodic forcing of ISNs. We show that ISNs can be excited over a range of frequencies and derive properties of the resonance peaks. In particular, we studied the phase-locked solutions, the torus solutions and the resonance peaks. More particularly, periodically forced ISNs respond with (possibly multi-stable) phase-locked activity whereas networks with sustained intrinsic oscillations respond more dynamically to periodic inputs with tori. Hence, the dynamics are surprisingly rich and phase effects alone do not adequately describe the network response. This strengthens the importance of phase-amplitude coupling as opposed to phase-phase coupling in providing multiple frequencies for multiplexing and routing information.

This work has been published in *Neural Computation* and is available as [29].

6.1.2. *A new twist for the simulation of hybrid systems using the true jump method*

Participant: Romain Veltz.

The use of stochastic models, in effect piecewise deterministic Markov processes (PDMP), has become increasingly popular especially for the modeling of chemical reactions and cell biophysics. Yet, exact simulation methods, for the simulation of these models in evolving environments, are limited by the need to find the next jumping time at each recursion of the algorithm. We report on a new general method to find this jumping time for the True Jump Method. It is based on an expression in terms of ordinary differential equations for which efficient numerical methods are available. As such, our new result makes it possible to study numerically stochastic models for which analytical formulas are not available thereby providing a way to approximate the state distribution for example. We conclude that the wide use of event detection schemes for the simulation of PDMPs should be strongly reconsidered. The only relevant remaining question being the efficiency of our method compared to the Fictitious Jump Method, question which is strongly case dependent.

This work is available as [55].

6.1.3. *On the effects on cortical spontaneous activity of the symmetries of the network of pinwheels in visual area V1*

Participants: Romain Veltz, Pascal Chossat, Olivier Faugeras.

This work challenges and extends earlier seminal work. We consider the problem of describing mathematically the spontaneous activity of V1 by combining several important experimental observations including (1) the organization of the visual cortex into a spatially periodic network of hypercolumns structured around pinwheels, (2) the difference between short-range and long-range intracortical connections, the first ones being rather isotropic and producing naturally doubly periodic patterns by Turing mechanisms, the second one being patchy, and (3) the fact that the Turing patterns spontaneously produced by the short-range connections and the network of pinwheels have similar periods. By analyzing the preferred orientation (PO) maps, we are able to classify all possible singular points (the pinwheels) as having symmetries described by a small subset of the

wallpaper groups. We then propose a description of the spontaneous activity of V1 using a classical voltage-based neural field model that features isotropic short-range connectivities modulated by non-isotropic long-range connectivities. A key observation is that, with only short-range connections and because the problem has full translational invariance in this case, a spontaneous doubly periodic pattern generates a 2-torus in a suitable functional space which persists as a flow-invariant manifold under small perturbations, for example when turning on the long-range connections. Through a complete analysis of the symmetries of the resulting neural field equation and motivated by a numerical investigation of the bifurcations of their solutions, we conclude that the branches of solutions which are stable over an extended range of parameters are those that correspond to patterns with an hexagonal (or nearly hexagonal) symmetry. The question of which patterns persist when turning on the long-range connections is answered by (1) analyzing the remaining symmetries on the perturbed torus and (2) combining this information with the Poincaré-Hopf theorem. We have developed a numerical implementation of the theory that has allowed us to produce the predicted patterns of activities, the planforms. In particular we generalize the contoured and non-contoured planforms predicted by previous authors.

This work has been published in Journal of Mathematical Neuroscience and is available as [27].

6.1.4. Biophysical reaction-diffusion model for stage II retinal waves and bifurcations analysis

Participants: Theodora Karvouniari, Bruno Cessac.

Retinal waves are spontaneous waves of spiking activity observed in the retina, during development only, playing a central role in shaping the visual system and retinal circuitry. Understanding how these waves are initiated and propagate in the retina could enable one to control, guide and predict them in the in vivo adult retina as inducing them is expected to reintroduce some plasticity in the retinal tissue and in the projections to the LGN. In this context, we propose a physiologically realistic reaction-diffusion model for the mechanisms of the emergence of stage II cholinergic retinal waves during development. We perform the bifurcation analysis when varying two biophysically relevant parameters, the conductances of calcium and potassium g_{Ca} , g_K respectively. The two main goals of our work are: firstly, reproduce the experimental recordings of developmental retinal waves by simulating our model and secondly, explore the different dynamical behaviours observed when varying these two parameters.

This work is available as [35].

6.1.5. Spatio-Temporal Linear Response of Spiking Neuronal Network Models

Participants: Rodrigo Cofré, Bruno Cessac.

We study the impact of a weak time-dependent external stimulus on the collective statistics of spiking responses in neuronal networks. We extend the current knowledge, assessing the impact over firing rates and cross correlations, to any higher order spatio-temporal correlation [1]. Our approach is based on Gibbs distributions (in a general setting considering non stationary dynamics and infinite memory) [2] and linear response theory. The linear response is written in terms of a correlation matrix, computed with respect to the spiking dynamics without stimulus. We give an example of application in a conductance based integrate-and fire model.

This work is available as [38].

6.1.6. Heteroclinic cycles in Hopfield networks

Participants: Pascal Chossat, Maciej Krupa.

It is widely believed that information is stored in the brain by means of the varying strength of synaptic connections between neurons. Stored patterns can be replayed upon the arrival of an appropriate stimulus. Hence, it is interesting to understand how an information pattern can be represented by the dynamics of the system. In this work, we consider a class of network neuron models, known as Hopfield networks, with a learning rule which consists of transforming an information string to a coupling pattern. Within this class of models, we study dynamic patterns, known as robust heteroclinic cycles, and establish a tight connection between their existence and the structure of the coupling.

This work has been published in Journal of Nonlinear Science and is available as [20].

6.2. Mean field approaches

6.2.1. *Confronting mean-field theories to measurements: a perspective from neuroscience*

Participant: Bruno Cessac.

Mean-field theories in neuroscience are usually understood as ways to bridge spatial and temporal scales by lumping together the activities of many single neurons, and then explaining or predicting the spatio-temporal variations of mesoscopic or macroscopic quantities measurable with current technologies: EEG, MEG, fMRI, optical imaging, etc. This is very much alike the situation in statistical physics where macroscopic quantities such as pressure, conductivity and so on are explained by the interactions between "microscopic" entities like atoms or molecules.

The situation in neuroscience is different however: the laws governing the microscopic dynamics in physics do not have the same structure as the laws governing neuronal dynamics; for example, interactions between neurons are not symmetric. Moreover, it is yet unclear what the relevant macroscopic quantities are in order to account for, say, visual perception. At the present stage of research, these quantities are considered to be what is measurable with currently available technologies, whereas better theories could reveal new types of phenomenological observables with a higher explanatory power.

We review mean-field methods coming from physics and their consequences on neuronal dynamics predictions.

This work is available as [30], [31], [32].

6.2.2. *A Formalism for Evaluating Analytically the Cross-Correlation Structure of a Firing-Rate Network Model*

Participants: Diego Fasoli, Olivier Faugeras, Stefano Panzeri.

We introduce a new formalism for evaluating analytically the cross-correlation structure of a finite size firing-rate network with recurrent connections. The analysis performs a first-order perturbative expansion of neural activity equations that include three different sources of randomness: the background noise of the membrane potentials, their initial conditions, and the distribution of the recurrent synaptic weights. This allows the analytical quantification of the relationship between anatomical and functional connectivity, i.e. of how the synaptic connections determine the statistical dependencies at any order among different neurons. The technique we develop is general, but for simplicity and clarity we demonstrate its efficacy by applying it to the case of synaptic connections described by regular graphs. The analytical equations obtained in this way reveal previously unknown behaviors of recurrent firing-rate networks, especially on how correlations are modified by the external input, by the finite size of the network, by the density of the anatomical connections and by correlation in sources of randomness. In particular, we show that a strong input can make the neurons almost independent, suggesting that functional connectivity does not depend only on the static anatomical connectivity, but also on the external inputs. Moreover we prove that in general it is not possible to find a mean-field description à la Sznitman of the network, if the anatomical connections are too sparse or our three sources of variability are correlated. To conclude, we show a very counterintuitive phenomenon, which we call stochastic synchronization, through which neurons become almost perfectly correlated even if the sources of randomness are independent. Due to its ability to quantify how activity of individual neurons and the correlation among them depends upon external inputs, the formalism introduced here can serve as a basis for exploring analytically the computational capability of population codes expressed by recurrent neural networks.

This work is available as [22].

6.2.3. *Asymptotic Description of Neural Networks with Correlated Synaptic Weights*

Participants: Olivier Faugeras, James Maclaurin.

We study the asymptotic law of a network of interacting neurons when the number of neurons becomes infinite. Given a completely connected network of neurons in which the synaptic weights are Gaussian correlated random variables, we describe the asymptotic law of the network when the number of neurons goes to infinity. We introduce the process-level empirical measure of the trajectories of the solutions to the equations of the finite network of neurons and the averaged law (with respect to the synaptic weights) of the trajectories of the solutions to the equations of the network of neurons. The main result of this article is that the image law through the empirical measure satisfies a large deviation principle with a good rate function which is shown to have a unique global minimum. Our analysis of the rate function allows us also to characterize the limit measure as the image of a stationary Gaussian measure defined on a transformed set of trajectories.

This work is available as [23].

6.2.4. Clarification and Complement to "Mean-Field Description and Propagation of Chaos in Networks of Hodgkin-Huxley and FitzHugh-Nagumo Neurons"

Participants: Mireille Bossy, Olivier Faugeras, Denis Talay.

In this work, we clarify the well-posedness of the limit equations to the mean-field N-neuron models proposed in [1] and we prove the associated propagation of chaos property. We also complete the modeling issue in [1] by discussing the well-posedness of the stochastic differential equations which govern the behavior of the ion channels and the amount of available neurotransmitters.

This work is available as [18].

6.3. Neural fields theory

6.3.1. ERRATUM: A Center Manifold Result for Delayed Neural Fields Equations

Participants: Romain Veltz, Olivier Faugeras.

Lemma C.1 in [95] is wrong. This lemma is used in the proof of the existence of a smooth center manifold, Theorem 4.4 in that paper. An additional assumption is required to prove this existence. We spell out this assumption, correct the proofs, and show that the assumption is satisfied for a large class of delay functions τ . We also weaken the general assumptions on τ .

This work has been published in SIAM J. Math. Anal. and is available as [28].

6.3.2. A general framework for stochastic traveling waves and patterns, with application to neural field equations

Participants: James Inglis, James Maclaurin.

In this work we present a general framework in which to rigorously study the effect of spatio-temporal noise on traveling waves and stationary patterns. In particular, the framework can incorporate versions of the stochastic neural field equation that may exhibit traveling fronts, pulses or stationary patterns. To do this, we first formulate a local SDE that describes the position of the stochastic wave up until a discontinuity time, at which point the position of the wave may jump. We then study the local stability of this stochastic front, obtaining a result that recovers a well-known deterministic result in the small-noise limit. We finish with a study of the long-time behavior of the stochastic wave.

This work has appeared in SIAM J. on Applied Dynamical Systems (SIADS) [49].

6.4. Slow-Fast Dynamics in Neural Models

6.4.1. From Canards of Folded Singularities to Torus Canards in a Forced van der Pol Equation

Participants: John Burke [Boston University, USA], Mathieu Desroches, Albert Granados [Technical University of Denmark, Lyngby, Denmark], Tasso Kaper [Boston University, USA], Maciej Krupa, Theodore Vo [Boston University, USA].

In this work, we study canard solutions of the forced van der Pol equation in the relaxation limit for low-, intermediate-, and high-frequency periodic forcing. A central numerical observation made herein is that there are two branches of canards in parameter space which extend across all positive forcing frequencies. In the low-frequency forcing regime, we demonstrate the existence of primary maximal canards induced by folded saddle nodes of type I and establish explicit formulas for the parameter values at which the primary maximal canards and their folds exist. Then, we turn to the intermediate- and high-frequency forcing regimes and show that the forced van der Pol possesses torus canards instead. These torus canards consist of long segments near families of attracting and repelling limit cycles of the fast system, in alternation. We also derive explicit formulas for the parameter values at which the maximal torus canards and their folds exist. Primary maximal canards and maximal torus canards correspond geometrically to the situation in which the persistent manifolds near the family of attracting limit cycles coincide to all orders with the persistent manifolds that lie near the family of repelling limit cycles. The formulas derived for the folds of maximal canards in all three frequency regimes turn out to be representations of a single formula in the appropriate parameter regimes, and this unification confirms the central numerical observation that the folds of the maximal canards created in the low-frequency regime continue directly into the folds of the maximal torus canards that exist in the intermediate- and high-frequency regimes. In addition, we study the secondary canards induced by the folded singularities in the low-frequency regime and find that the fold curves of the secondary canards turn around in the intermediate-frequency regime, instead of continuing into the high-frequency regime. Also, we identify the mechanism responsible for this turning. Finally, we show that the forced van der Pol equation is a normal form-type equation for a class of single-frequency periodically driven slow/fast systems with two fast variables and one slow variable which possess a non-degenerate fold of limit cycles. The analytic techniques used herein rely on geometric desingularisation, invariant manifold theory, Melnikov theory, and normal form methods. The numerical methods used herein were developed in Desroches et al. (SIAM J Appl Dyn Syst 7:1131–1162, 2008, Nonlinearity 23:739–765 2010).

This work has been published in J. Nonlinear Sci. and is available as [19].

6.4.2. *Extending the zero-derivative principle for slow-fast dynamical systems*

Participants: Eric Benoît [Université de La Rochelle, France], Morten Brøns [Technical University of Denmark, Lyngby, Denmark], Mathieu Desroches, Maciej Krupa.

Slow-fast systems often possess slow manifolds, that is invariant or locally invariant sub-manifolds on which the dynamics evolves on the slow time scale. For systems with explicit timescale separation, the existence of slow manifolds is due to Fenichel theory, and asymptotic expansions of such manifolds are easily obtained. In this work, we discuss methods of approximating slow manifolds using the so-called zero-derivative principle. We demonstrate several test functions that work for systems with explicit time scale separation including ones that can be generalized to systems without explicit timescale separation. We also discuss the possible spurious solutions, known as ghosts, as well as treat the Templator system as an example.

This work has been published in ZAMP and is available as [17].

6.4.3. *Canards, folded nodes and mixed-mode oscillations in piecewise-linear slow-fast systems*

Participants: Mathieu Desroches, Antoni Guillamon [Polytechnic University of Catalunya, Barcelona, Spain], Enrique Ponce [University of Sevilla, Spain], Rafel Prohens [University of the Balearic Islands, Palma, Spain], Serafim Rodrigues [Plymouth University, UK], Antonio Teruel [University of the Balearic Islands, Palma, Spain].

Canard-induced phenomena have been extensively studied in the last three decades, both from the mathematical and from the application viewpoints. Canards in slow-fast systems with (at least) two slow variables, especially near folded-node singularities, give an essential generating mechanism for Mixed-Mode oscillations (MMOs) in the framework of smooth multiple timescale systems. There is a wealth of literature on such slow-fast dynamical systems and many models displaying canard-induced MMOs, in particular in neuroscience. In parallel, since the late 1990s several papers have shown that the canard phenomenon can be faithfully reproduced with piecewise-linear (PWL) systems in two dimensions although very few results are available in the

three-dimensional case. This work aims to bridge this gap by analyzing canonical PWL systems that display folded singularities, primary and secondary canards, with a similar control of the maximal winding number as in the smooth case. We also show that the singular phase portraits are compatible in both frameworks. Finally, we show on an example how to construct a (linear) global return and obtain robust PWL MMOs.

This work has been accepted for publication in SIAM Review and is available as [46].

6.4.4. *Canard solutions in planar piecewise linear systems with three zones*

Participants: Soledad Fernández-García [Inria Paris-Rocquencourt, France], Mathieu Desroches, Maciej Krupa, Antonio Teruel [University of the Balearic Islands, Palma, Spain].

In this work, we analyze the existence and stability of canard solutions in a class of planar piecewise linear systems with three zones, using a singular perturbation theory approach. To this aim, we follow the analysis of the classical canard phenomenon in smooth planar slow-fast systems and adapt it to the piecewise-linear framework. We first prove the existence of an intersection between repelling and attracting slow manifolds, which defines a maximal canard, in a non-generic system of the class having a continuum of periodic orbits. Then, we perturb this situation and prove the persistence of the maximal canard solution, as well as the existence of a family of canard limit cycles in this class of systems. Similarities and differences between the piecewise linear case and the smooth one are highlighted.

This work has been published *Dynam. Syst.* and is available as [24].

6.4.5. *Spike-adding mechanism in parabolic bursters: the role of folded-saddle canards*

Participants: Mathieu Desroches, Maciej Krupa, Serafim Rodrigues [Plymouth University, UK].

The present work develops a new approach to studying parabolic bursting, and also proposes a novel four-dimensional canonical and polynomial-based parabolic burster. In addition to this new polynomial system, we also consider the conductance-based model of the Aplysia R15 neuron known as Plant's model, and a reduction of this prototypical biophysical parabolic burster to three variables, including one phase variable, namely Rinzel's theta model. Revisiting these models from the perspective of slow-fast dynamics reveals that the number of spikes per burst may vary upon parameter changes, however the spike-adding process occurs in a brutal (explosive) fashion that involves special solutions called canards. This spike-adding canard explosion phenomenon is analysed by using tools from geometric singular perturbation theory in tandem with numerical bifurcation techniques. We find that the bifurcation structure persists across both parabolic bursters, that is, spikes within the burst are incremented via the crossing of an excitability threshold given by a particular type of canard orbit, namely the strong canard of a folded-saddle singularity. Using these findings, we construct a new polynomial approximation of Plant's model, which retains all the key elements for parabolic bursting, including the canard mediated spike-adding transitions. Finally, we briefly investigate the presence of spike-adding via canards in planar phase models of parabolic bursting, namely the theta model by Ermentrout and Kopell.

This work has been submitted for publication and is available as [47].

6.4.6. *Canards and spike-adding transitions in a minimal piecewise-linear Hindmarsh-Rose square-wave burster*

Participants: Mathieu Desroches, Soledad Fernández-García [Inria Paris-Rocquencourt, France], Maciej Krupa.

We construct a piecewise-linear (PWL) approximation of the Hindmarsh-Rose (HR) neurone model that is minimal, in the sense that the vector field has the least number of pieces, in order to reproduce all the dynamics present in the original HR model with the classical parameter values. This includes spiking, square-wave bursting, and also special trajectories called canards, which possess long repelling segments and organise the transition between stable bursting patterns with n and $n + 1$ spikes. This is the spike-adding canard explosion. We propose a first approximation of the smooth bursting model, using a continuous PWL system, and show that its fast subsystem cannot possess a homoclinic bifurcation, which is necessary to obtain proper square-wave bursting. We then relax the assumption of continuity of the vector field across all zones and show that

we can obtain a homoclinic bifurcation in the fast subsystem. We use the recently developed canard theory for PWL systems in order to reproduce the spike-adding canard explosion feature of the HR model as studied, e.g., in [66].

This work has been submitted for publication and is available as [45].

6.4.7. Ducks in space

Participants: Daniele Avitabile [University of Nottingham, UK], Mathieu Desroches, Edgar Knobloch [University of California at Berkeley, USA], Maciej Krupa.

A subcritical pattern-forming system with nonlinear advection in a bounded domain is recast as a slow-fast system in space and studied using a combination of geometric singular perturbation theory and numerical continuation. Two types of solutions describing the possible location of stationary fronts are identified, one of which is present for all values of the bifurcation parameter while the other is present for zero or sufficiently small inlet boundary conditions but only when the bifurcation parameter is large enough. For slightly larger inlet boundary condition a continuous transition from one type to the other takes place as the bifurcation parameter increases. The origin of the two solution types is traced to the onset of convective and absolute instability on the real line. The role of canard trajectories in the transitions between these states is clarified and the stability properties of the resulting spatial structures are determined. Front location in the convective regime is highly sensitive to the upstream boundary condition and its dependence on this boundary condition is studied using a combination of numerical continuation and Monte Carlo simulations of the partial differential equation. Statistical properties of the system subjected to random or stochastic boundary conditions are interpreted using the deterministic slow-fast spatial-dynamical system.

This work has been submitted for publication and is available as [43].

6.5. Spike Train statistics

6.5.1. Statistical models for spike trains analysis in the retina.

Participant: Bruno Cessac.

Recent advances in multi-electrodes array acquisition have made it possible to record the activity of up to several hundreds of neurons at the same time and to register their collective activity (spike trains). For the retina, this opens up new perspectives in understanding how retinal structure and ganglion cells encode information about a visual scene and what is transmitted to the brain. Especially, two paradigms can be confronted: in the first one, ganglion cells encode information independently of each others; in the second one non linear dynamics and connectivity contribute to produce a population coding where spatio-temporal correlations, although weak, play a significant role in spike coding. Confronting these two paradigms can be done at an experimental and at a theoretical level. On experimental grounds, new methods to analyse the role of weak correlations in spike train statistics are required. On theoretical grounds, mathematical results have been established, in neuronal models, showing how non linear dynamics and connectivity contribute to produce a correlated spike response to stimuli. In the context of the ANR KEOPS project, we have been working on these two aspects and we present our main results.

This work is available as [33].

6.5.2. Spectral dimension reduction on parametric models for spike train statistics

Participants: Cesar Ravello, Ruben Herzog, Bruno Cessac, Maria-Jose Escobar, Adrian Palacios.

It has been shown that the neurons of visual system present correlated activity in response to different stimuli. The role of these correlations is an unresolved subject. These correlations vary according to the stimulus, specially with natural images. To uncover the role of these correlation and characterize the population code, it is necessary to measure the simultaneous activity of large neural populations. This has been achieved thanks to the advent of Multi-Electrode Array technology, opening up a way to better characterize how the brain encodes information in the concerted activity of neurons. In parallel, powerful statistical tools have been developed to accurately characterize spatio-temporal correlations between neurons. Methods based on *Maximum Entropy Principle*, where statistical entropy is maximized under a set of constraints corresponding to specific assumptions on the relevant statistical quantities, have been proved successfully, specially when they consider *spatiotemporal* correlations. They are although limited by (i) **the assumption of stationarity**, (ii) **the many possible choice of constraints**, and (iii) **the huge number of free parameters**. We present our results on these two aspects obtained in the context of **ANR KEOPS**.

This work is available as [54].

6.6. Visual Neuroscience

6.6.1. Shifting stimulus for faster receptive fields estimation of ensembles of neurons

Participants: Bruno Cessac, Matthias Hennig [University of Edinburg, UK], Gerrit Hilgen [Institute of Neuroscience, Medical School, Newcastle University, Newcastle, UK], Pierre Kornprobst, Daniela Pamplona, Sahar Pirmoradian [University of Edinburg, UK], Evelyne Sernagor [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK].

The Spike Triggered Average (STA) is a classical technique to find a discrete approximation of the Receptive Fields (RFs) of sensory neurons [63], a required analysis in most experimental studies. One important parameter of the STA is the spatial resolution of the estimation, corresponding to the size of the blocks of the checkerboard stimulus images. In general, it is experimentally fixed to reach a compromise: If too small, neuronal responses might be too weak thus leading to RF with low Signal-to-Noise-Ratio; on the contrary, if too large, small RF will be lost, or not described with enough details, because of the coarse approximation. Other solutions were proposed consisting in starting from a small block size and updating it following the neuron response in a closed-loop to increase its response [70], [78], [77]. However, these solutions were designed for single cells and cannot be applied to simultaneous recordings of ensembles of neurons (since each RF has its own size and preferred stimulus).

To solve this problem, we introduced a modified checkerboard stimulus where blocks are shifted randomly in space at fixed time steps. This idea is inspired from super-resolution techniques developed in image processing [84]. The main interest is that the block size can be large, enabling strong responses, while the resolution can be finer since it depends on the shift minimum size. In [52], we show that the STA remains an unbiased RF estimator and, using simulated spike trains from an ensemble of Linear Nonlinear Poisson cascade neurons, it was predicted that this approach improves RF estimation over the neuron ensemble, in terms of resolution and convergence. In [53], we test these predictions experimentally on the RFs estimation of 8460 ganglion cells from two mouse retinas, using recordings performed with a large scale high-density multielectrode array. We compare RFs obtained using (i) the classical checkerboard stimulus with block size of $160\mu\text{m}$ and (ii) our checkerboard stimulus with block size of $160\mu\text{m}$ and arbitrary shifts of $40\mu\text{m}$ in x - and y -directions. Results show how spatial resolution can be improved and that our approach allows to recover 51% of the mapped RFs at a resolution of $40\mu\text{m}$, while in the classical case, 41% of the RFs could be found at a resolution of only $160\mu\text{m}$. Thus, our approach improves not only the quality of the RF estimation but also the amount of successfully mapped RFs in neural ensembles.

This work was presented in [52], [53] and it is being used in current experimental protocols by E. Sernagor (Newcastle University), partner of the EC IP project FP7-ICT-2011-9 no. 600847 (RENVISION).

6.6.2. Using neural mechanisms underlying motion analysis for optical flow estimation

Participants: Manuela Chessa [University of Genoa, DIBRIS, Italy], Pierre Kornprobst, Guillaume S. Masson [Institut de Neurosciences de la Timone, Team InVibe], Kartheek Medathati, Fabio Solari [University of Genoa, DIBRIS, Italy].

We explore how motion information, also called optical flow, is estimated from natural moving sequences. Owing to application potential, optical flow estimation has been studied extensively by computer vision. On the other hand the neural mechanisms underlying motion analysis in the visual cortex have been extensively studied almost with little interaction with computer vision community resulting in few mathematical models. Even though there was some early interaction among the two communities for example, methods by Heeger et.al, Sejnowski et. al, comparatively little work has been done in terms of examining or extending the mathematical models proposed in biology in terms of their engineering efficacy on modern optical flow estimation datasets.

Pursuing this idea, in [26], we proposed a neural model inspired from the ones presented in [87], [86] which are popular models of primate velocity encoding. We started from a classical V1-MT feedforward architecture. We modeled V1 cells by motion energy (based on spatio-temporal filtering), and MT pattern cells (by pooling V1 cell responses). The efficacy of this architecture and its inherent limitations in the case of real videos were not known. To answer this question, we proposed a velocity space sampling of MT neurones (using a decoding scheme to obtain the local velocity from their activity) coupled with a multi-scale approach. After this, we explored the performance of our model on the Middlebury dataset. To the best of our knowledge, this is the only neural model in this dataset. The results were promising and suggested several possible improvements, in particular to better deal with discontinuities. An extension was proposed in [40].

We also focused on the decoding the motion energies which is of natural interest for developing biologically inspired computer vision algorithms for dense optical flow estimation. In [37], we addressed this problem by evaluating four strategies for motion decoding: intersection of constraints, maximum likelihood, linear regression on MT responses and neural network based regression using multi scale-features. We characterized the performances and the current limitations of the different strategies, in terms of recovering dense flow estimation using Middlebury benchmark dataset widely used in computer vision, and we highlight key aspects for future developments.

This work was partially funded by the EC IP project FP7-ICT-388 2011-8 no. 318723 (MatheMACS).

6.6.3. Bio-Inspired Computer Vision: Towards a Synergistic Approach of Artificial and Biological Vision

Participants: Pierre Kornprobst, Guillaume S. Masson [Institut de Neurosciences de la Timone, Team InVibe], Kartheek Medathati, Heiko Neumann [Ulm University, Germany].

Studies in biological vision have always been a great source of inspiration for design of computer vision algorithms. In the past, several successful methods were designed with varying degrees of correspondence with biological vision studies, ranging from purely functional inspiration to methods that utilise models that were primarily developed for explaining biological observations. Even though it seems well recognised that computational models of biological vision can help in design of computer vision algorithms, it is a non-trivial exercise for a computer vision researcher to mine relevant information from biological vision literature as very few studies in biology are organised at a task level.

In [42], we aim to bridge this gap by providing a computer vision task centric presentation of models primarily originating in biological vision studies. Not only we revisit some of the main features of biological vision and discuss the foundations of existing computational studies modelling biological vision, but also consider three classical computer vision tasks from a biological perspective: image sensing, segmentation and optical flow. Using this task-centric approach, we discuss well-known biological functional principles and compare them with approaches taken by computer vision. Based on this comparative analysis of computer and biological vision, we present some recent models in biological vision and highlight a few models that we think are promising for future investigations in computer vision. To this extent, this paper provides new insights and a starting point for investigators interested in the design of biology-based computer vision algorithms and pave a way for much needed interaction between the two communities leading to the development of synergistic models of artificial and biological vision.

[42] is under review. This work was partially funded by the EC IP project FP7-ICT-388 2011-8 no. 318723 (MatheMACS).

NEUROSYS Project-Team

7. New Results

7.1. From the Microscopic to the Mesoscopic Scale

Participants: Laure Buhry, Axel Hutt, Francesco Giovannini, Jean-Baptiste Schneider
In collaboration with LieJune Shiau (University of Houston)

7.1.1. Memory and Anaesthesia

7.1.1.1. Hippocampal Memory Networks

To improve our understanding of the effects of anaesthesia on the neural correlates of memory, we focussed on how anaesthetics disrupt the interaction between the hippocampus and the cerebral cortex. As a first step towards this objective Francesco Giovannini modelled a hippocampal pyramidal neuron using the Hodgkin-Huxley model capable of exhibiting long-lasting persistent firing activity when subject to a strong transient stimulus [16]. This behaviour is underlay by an intrinsic membrane current activated by the increase of intracellular calcium ions, following the discharge of an action potential by the neuron, in accord with that displayed in neural recordings of hippocampal slice preparations. Connecting these persistent firing neurons in a network comprising strong local excitation yields a wide range of behaviours depending on the interaction between CAN and synaptic currents. Indeed, the network model is capable of displaying rhythmic behaviour in the form of short synchronised bursts with intra-burst frequencies of 20 – 40 Hz and inter-burst frequencies of 3 Hz. Furthermore, coupling CAN-equipped pyramidal neurons with a population of fast-spiking inhibitory interneurons yields emerging synchronous activity whose frequency is modulated by the strength of this coupling. These results hint towards a possible mechanism for the generation of memory-related oscillatory activity in the hippocampus.

7.1.1.2. Anaesthetic Effects on Hippocampal Oscillations

We investigated the effects of propofol-mediated tonic inhibition on the synchronous activity elicited in a network of hippocampal inhibitory interneurons. This work was conducted in collaboration with Jean-Baptiste Schneider, as part of his 2-month internship. We studied the effect of propofol-induced tonic inhibition on the oscillations elicited in a network of hippocampal Hodgkin-Huxley gamma-aminobutyric acid (GABA) interneurons by studying the action of propofol on extrasynaptic GABAergic receptors. Our results [15] show that increasing doses of propofol reduce the overall network activity and slow down its oscillations until a critical value at which the synchronisation increases abruptly at values of twice the synchronisation displayed in the absence of tonic inhibition, and the mean firing rate increases. This emergence of synchronous activity mediated by anaesthetic perfusion point towards a possible mechanism for the emergence of paradoxical excitation under general anaesthesia.

In this context, Laure Buhry works with LieJune Shiau (University of Houston) on a better understanding of the models used by the community of computational neuroscientists. The goal is to show in which extent models are comparable or interchangeable. We focus on the comparison of oscillatory mechanisms of neuronal populations in different spiking models, especially in the Hodgkin-Huxley and the adaptive exponential integrate-and-fire model (AdEx). Especially, we have shown that a same number of synaptic connection per neuron is necessary to elicit synchronization in inhibitory neural networks of adaptive exponential integrate and fire neurons as in networks of Hodgkin-Huxley neurons. We have also conducted an extensive study regarding the effects of the different parameters of the AdEx model on the synchronization mechanisms in inhibitory neural networks, particularly in the context of gamma oscillations. A manuscript will be submitted soon to the Journal of Computational Neuroscience.

7.1.1.3. Noise Effects on Neural Rhythms

We have continued working on the effect of additive noise on neural oscillations and have shown that additive noise modulates the frequency of self-sustained neural rhythms [3].

7.2. From the Mesoscopic to the Macroscopic Scale

Participants: Laurent Bougrain, Axel Hutt, Pedro Garcia-Rodriguez, Eric Nichols, Guillaume Serrière, Tamara Tomic, Mariia Fedotenkova, Meysam Hashemi, Benjamin le Golvan, Cecilia Lindig-Leon, Sébastien Rimbart.

7.2.1. Level of Consciousness

7.2.1.1. Spatio-temporal Dynamics in Neural Fields

Neural fields serve as a model for experimental macroscopic activity. We have developed a numerical simulator NeuralFieldSimulator [21]. In addition, we have worked out a neural neural field model that exhibits a sequence of metastable activity states as observed in experimental data [4].

7.2.1.2. Synchronisation in Local Field Potentials under Anaesthesia

We have applied advanced data analysis techniques based on wavelet analysis to detect instantaneous partial synchronisation in experimental data [5].

7.2.1.3. Statistical Frequency-dependent Analysis by Recurrence Plots

Participants : Axel Hutt, Mariia Fedotenkova, Tamara Tomic

In collaboration with Flavio Frohlich, Peter Beim Graben and Kristin K. Sellers

For decades, research in neuroscience has supported the hypothesis that brain dynamics exhibits recurrent metastable states connected by transients, which together encode fundamental neural information processing. To understand the system's dynamics it is important to detect such recurrence domains, but it is challenging to extract them from experimental neuroscience datasets due to the large trial-to-trial variability. We proposed a methodology to extract recurrent metastable states in univariate time series by transforming datasets into their time-frequency representations and computing recurrence plots based on instantaneous spectral power values in various frequency bands [6]. Additionally, a new statistical inference analysis compares different trial recurrence plots with corresponding surrogates to obtain statistically significant recurrent structures. This combination of methods is validated by applying it to two artificial datasets. In a final study of visually-evoked Local Field Potentials in partially anesthetized ferrets, the methodology is able to reveal recurrence structures of neural responses with trial-to-trial variability. Focusing on different frequency bands, the delta-band activity is much less recurrent than alpha-band activity. Moreover, alpha-activity is susceptible to pre-stimuli, while delta-activity is much less sensitive to pre-stimuli. This difference in recurrence structures in different frequency bands indicates diverse underlying information processing steps in the brain.

7.2.2. Motor System

Participants: Laurent Bougrain, Axel Hutt, Benjamin le Golvan, Cecilia Lindig-Leon, Sébastien Rimbart, Guillaume Serrière

7.2.2.1. Motor Patterns during General Anesthesia

Participants: Laurent Bougrain, Axel Hutt, Cecilia Lindig-Leon, Sébastien Rimbart, Guillaume Serrière

The dosage of the anesthetic agent is tricky: too low, it does not achieve a sufficient loss of consciousness and may lead to a partial memorization during surgery and a post-operative trauma; too strong, it is dangerous for people with respiratory or heart problems. To better monitor the effect of the current dosage, we propose to study the dynamics of the motor brain activity during anesthesia. The relationship between motor brain activity and anesthesia is not intensively studied. Yet even if no physical movement by the patient is visually detectable, an electroencephalographic analysis of brain activity in motor areas may reveal an intention movement. This information is important because it demonstrates that the patient is conscious. We started to define a clinical protocol in collaboration with anesthesiologists of the hospital in Nancy to investigate its possibility. To reduce the duration of the protocol, we studied the minimum duration of a motor imagery to allow its detection from EEG recordings [23]. A large number of Brain-Computer Interfaces (BCIs) are based on the detection of motor imagery related features in the electroencephalographic signal. In most BCI experimental paradigms, subjects realize continuous motor imagery, i.e. a prolonged intention of movement, during a time window of a few seconds. Then, the system detects the movement based on the event-related desynchronization (ERD) and the event-related synchronization (ERS) principles. We studied if a discrete motor imagery, corresponding to a single short motor imagery, would allow a better detection of ERD and ERS patterns than a continuous motor

imagery. Indeed, the results of experiments involving 11 healthy subjects suggest that a continuous motor imagery generates a later ERS as well as a more variable and less detectable ERD than discrete motor imagery [11]. This finding suggests an improved experimental paradigm. We deeper investigated the amplitude and latency of EEG Beta activity during real movements, discrete and continuous motor imageries [22].

7.2.2.2. *Motor Patterns during Combined Movements*

Participants: Laurent Bougrain, Cecilia Lindig-Leon

Imaginary motor tasks cause brain oscillations that can be detected through the analysis of electroencephalographic (EEG) recordings. We studied whether or not the characteristics of the brain activity induced by the combined motor imagery (MI) of both hands can be assumed as the superposition of the activity generated during simple hand MIs. After analyzing the sensorimotor rhythms in EEG signals of five healthy subjects, results show that the imagination of both hands movement generates in each brain hemisphere similar activity as the one produced by each simple hand MI in the contralateral side [8]. Furthermore, during simple hand MIs, brain activity over the ipsilateral hemisphere presents similar characteristics as those observed during the rest condition. Thus, it is shown that the proposed scheme is valid and promising for brain-computer interfaces (BCI) control, allowing to easily detect patterns induced by combined MIs. Based on these results, we proposed a new method to extend the classic Common Spatial Pattern (CSP) algorithm to a multi-class approach which analyses both brain hemispheres separately to solve, together with a stepwise classification strategy, a multi-label BCI problem. After testing the proposed approach over the EEG signals of six healthy subjects performing a four-class multi-label task involving simple and combined hand MIs together with the rest condition, results show that this technique is plausible for BCI control [7]. In terms of accuracy, it outperforms the classical one-vs-one approach by 20% and has the same performance as the one-vs-all method. Nevertheless, to solve a multi-label classification problem involving k classes, the proposed method requires only $\log_2(k)$ classifiers, whereas the one-vs-one method uses $k(k-1)/2$ classifiers and the one-vs-all k classifiers, thereby the new approach simplifies the classification task and seems promising for solving multi-label problems involving numerous classes.

7.2.2.3. *On-line Detection of the End of Motor Imageries*

Participants : Cécilia Lindig-León, Laurent Bougrain and Sébastien Rimbart

Limb movement execution or imagination induce sensorimotor rhythms that can be detected in electroencephalographic (EEG) recordings. We presented the interest of considering not only the beta frequency band but also the alpha band to detect the elicited EEG rebound, i.e. the increasing of oscillatory power synchronization, at the end of motor imageries [9], [19]. This phenomenon can be stronger over the alpha than the beta band and it is experimentally demonstrated [9] that the analysis on the alpha band improves the detection of the end of motor imageries. Moreover a variant method to compute the oscillatory power without referring to a baseline period is proposed; such capacity is useful for self-paced BCI control.

7.2.3. *Pain under General Anaesthesia*

7.2.3.1. *Detection of EEG-signal Features for Pain under General Anaesthesia*

Participants : Axel Hutt, Mariia Fedotenkova

In collaboration with Peter Beim Graben and James W. Sleigh

Nowadays, surgical operations are impossible to imagine without general anaesthesia, which involves loss of consciousness, immobility, amnesia and analgesia. Understanding mechanisms underlying each of these effects guarantees well-controlled medical treatment. Our work focuses on analgesia effect of general anaesthesia, more specifically, on patients reaction on nociception stimuli. The study was conducted on dataset consisting of 230 EEG signals: pre- and post-incisional recordings for 115 patients, who received desflurane and propofol. Initial analysis was performed by power spectral analysis, which is a widespread approach in signal processing. Power spectral information was described by fitting the background activity and measuring power contained in delta and alpha bands according to power of background activity. The fact that power spectrum of background activity decays as frequency increasing is well known and thoroughly studied. Here, traditional $1/f^\alpha$ behaviour of the decay was replaced by a Lorentzian model to describe the power spectrum of background activity. Due to observed non-stationary nature of EEG signals spectral analysis does not suffice

to reveal significant changes between two states. A further improvement was done by expanding spectra with time information. To obtain time-frequency representations of the signals conventional spectrograms were used as well as a spectrogram reassignment technique. The latter allows to ameliorate readability of a spectrogram by reassigning energy contained in spectrogram to more precise positions. Subsequently, obtained spectrograms were used in recurrence analysis and its quantification by complexity measure. Recurrence analysis allows to describe and visualise dynamics of a system and discover structural patterns contained in the data. Structure of each recurrence plot is characterised by Lempel–Ziv complexity measure [5], which shows a difference between pre- and post-incision [13].

NUMED Project-Team (section vide)

PARIETAL Project-Team

7. New Results

7.1. Semi-Supervised Factored Logistic Regression for High-Dimensional Neuroimaging Data

Imaging neuroscience links human behavior to aspects of brain biology in ever-increasing datasets. Existing neuroimaging methods typically perform either discovery of unknown neural structure or testing of neural structure associated with mental tasks. However, testing hypotheses on the neural correlates underlying larger sets of mental tasks necessitates adequate representations for the observations. We therefore propose to blend representation modelling and task classification into a unified statistical learning problem. A multinomial logistic regression is introduced that is constrained by factored coefficients and coupled with an auto-encoder. We show that this approach yields more accurate and interpretable neural models of psychological tasks in a reference dataset, as well as better generalization to other datasets.

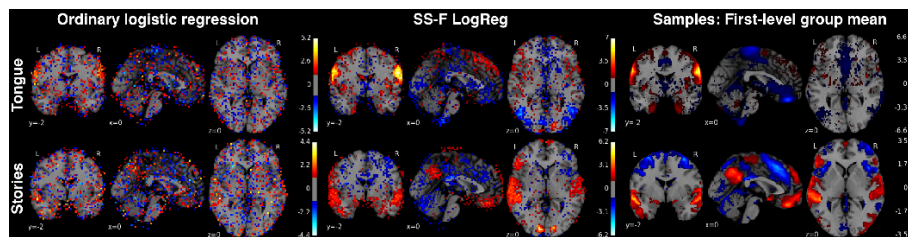


Figure 3. Classification weight maps. The voxel predictors corresponding to 2 exemplary (of 18 total) psychological tasks (rows) from the Human Connectome Project dataset. Left column: multinomial logistic regression (same implementation but without bottleneck or autoencoder), middle column: Semi-Supervised Factored Logistic Regression (SSFLogReg), right column: voxel-wise average across all samples of whole-brain activity maps from each task. SSFLogReg puts higher absolute weights on relevant structure, lowers ones on irrelevant structure, and yields BOLD-typical local contiguity (without enforcing an explicit spatial prior). More information can be found in [50].

More information can be found in [50].

7.2. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain

Here we present NeuroVault — a web based repository that allows researchers to store, share, visualize, and decode statistical maps of the human brain. NeuroVault is easy to use and employs modern web technologies to provide informative visualization of data without the need to install additional software. In addition, it leverages the power of the Neurosynth database to provide cognitive decoding of deposited maps. The data are exposed through a public REST API enabling other services and tools to take advantage of it. NeuroVault is a new resource for researchers interested in conducting meta- and coactivation analyses.

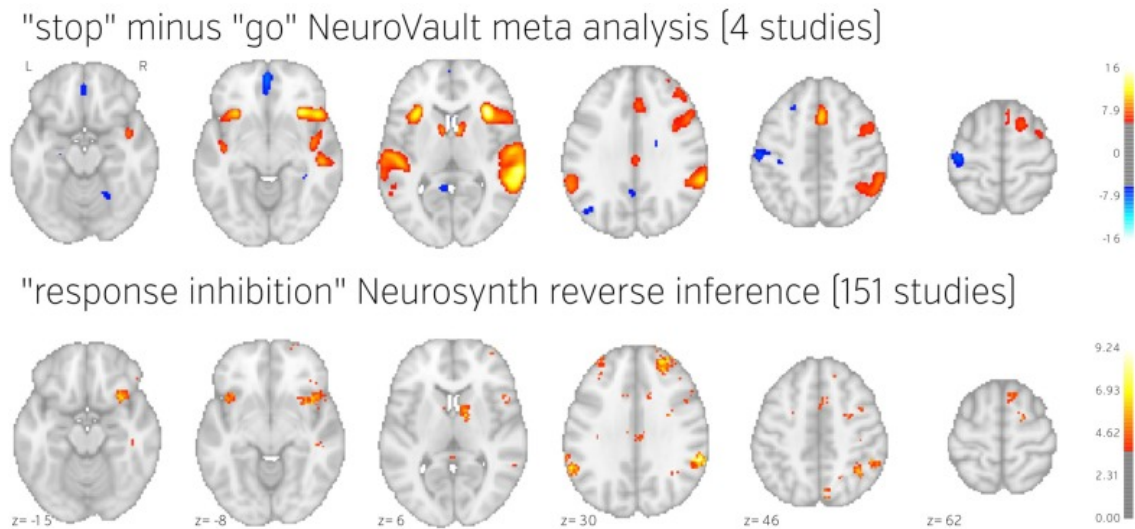


Figure 4.

Comparison of image based and coordinate based meta analysis of response inhibition. Meta analysis based on unthresholded statistical maps obtained from NeuroVault (top row) managed to recover the pattern of activation obtained using traditional methods despite including much fewer studies. NeuroVault map has been thresholded at $z = 6$, response inhibition map has been thresholded at $z = 1.77$ (the threshold values were chosen for visualization purposes only, but both are statistically significant at $p < 0.05$). Unthresholded versions of these maps are available at <http://neurovault.org/collections/439/>

More information can be found in [18] and [17].

7.3. FAASTA: A fast solver for total-variation regularization of ill-conditioned problems with application to brain imaging

The total variation (TV) penalty, as many other analysis-sparsity problems, does not lead to separable factors or a proximal operator with a closed-form expression, such as soft thresholding for the ℓ_1 penalty. As a result, in a variational formulation of an inverse problem or statistical learning estimation, it leads to challenging non-smooth optimization problems that are often solved with elaborate single-step first-order methods. When the data-fit term arises from empirical measurements, as in brain imaging, it is often very ill-conditioned and without simple structure. In this situation, in proximal splitting methods, the computation cost of the gradient step can easily dominate each iteration. Thus it is beneficial to minimize the number of gradient steps. We present fAASTA, a variant of FISTA, that relies on an internal solver for the TV proximal operator, and refines its tolerance to balance computational cost of the gradient and the proximal steps. We give benchmarks and illustrations on “brain decoding”: recovering brain maps from noisy measurements to predict observed behavior. The algorithm as well as the empirical study of convergence speed are valuable for any non-exact proximal operator, in particular analysis-sparsity problems.

Convergence of currently available optimization algorithms, for 3 scenarios, with weak, medium and strong regularization, where medium regularization corresponds to the value chosen by cross-validation. These are log-log plots with the 0 defined as the lowest energy value reached across all algorithms.

More information can be found in [47].

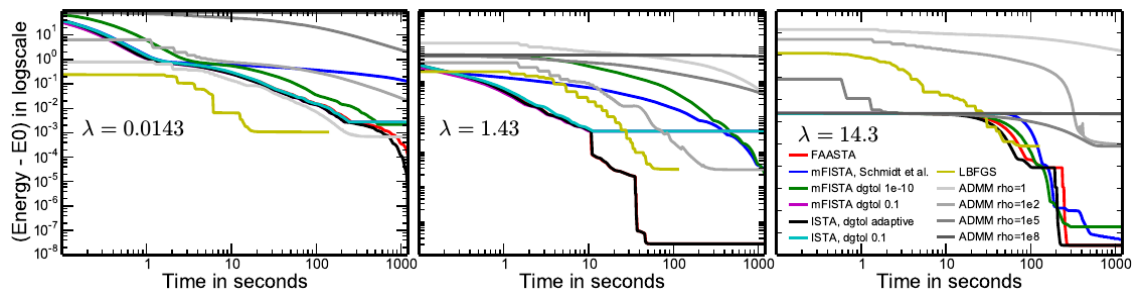


Figure 5.

7.4. Bootstrapped Permutation Test for Multiresponse Inference on Brain Behavior Associations.

Despite that diagnosis of neurological disorders commonly involves a collection of behavioral assessments, most neuroimaging studies investigating the associations between brain and behavior largely analyze each behavioral measure in isolation. To jointly model multiple behavioral scores, sparse multi-response regression (SMR) is often used. However, directly applying SMR without statistically controlling for false positives could result in many spurious findings. For models, such as SMR, where the distribution of the model parameters is unknown, permutation test and stability selection are typically used to control for false positives. In this paper, we present another technique for inferring statistically significant features from models with unknown parameter distribution. We refer to this technique as bootstrapped permutation test (BPT), which uses Studentized statistics to exploit the intuition that the variability in parameter estimates associated with relevant features would likely be higher with responses permuted. On synthetic data, we show that BPT provides higher sensitivity in identifying relevant features from the SMR model than permutation test and stability selection, while retaining strong control on the false positive rate. We further apply BPT to study the associations between brain connectivity estimated from pseudo-rest fMRI data of 1139 fourteen year olds and behavioral measures related to ADHD. Significant connections are found between brain networks known to be implicated in the behavioral tasks involved. Moreover, we validate the identified connections by fitting a regression model on pseudo-rest data with only those connections and applying this model on resting state fMRI data of 337 left out subjects to predict their behavioral scores. The predicted scores are shown to significantly correlate with the actual scores of the subjects, hence verifying the behavioral relevance of the found connections.

Real data results: Statistically significant connectivity differences between populations (a) Significant network connections found on pseudo-rest fMRI data. (b) Pearson's correlation between predicted and actual scores with p-values noted. Each set of three bars (top to bottom) correspond to spatial working memory strategy, spatial working memory between errors, and rapid visual information processing accuracy scores. Significance is declared at $p < 0.05$.

More information can be found in [43].

7.5. Total Variation meets Sparsity: statistical learning with segmenting penalties

Prediction from medical images is a valuable aid to diagnosis. For instance, anatomical MR images can reveal certain disease conditions, while their functional counterparts can predict neuropsychiatric phenotypes. However, a physician will not rely on predictions by black-box models: understanding the anatomical or functional features that underpin decision is critical. Generally, the weight vectors of classifiers are not

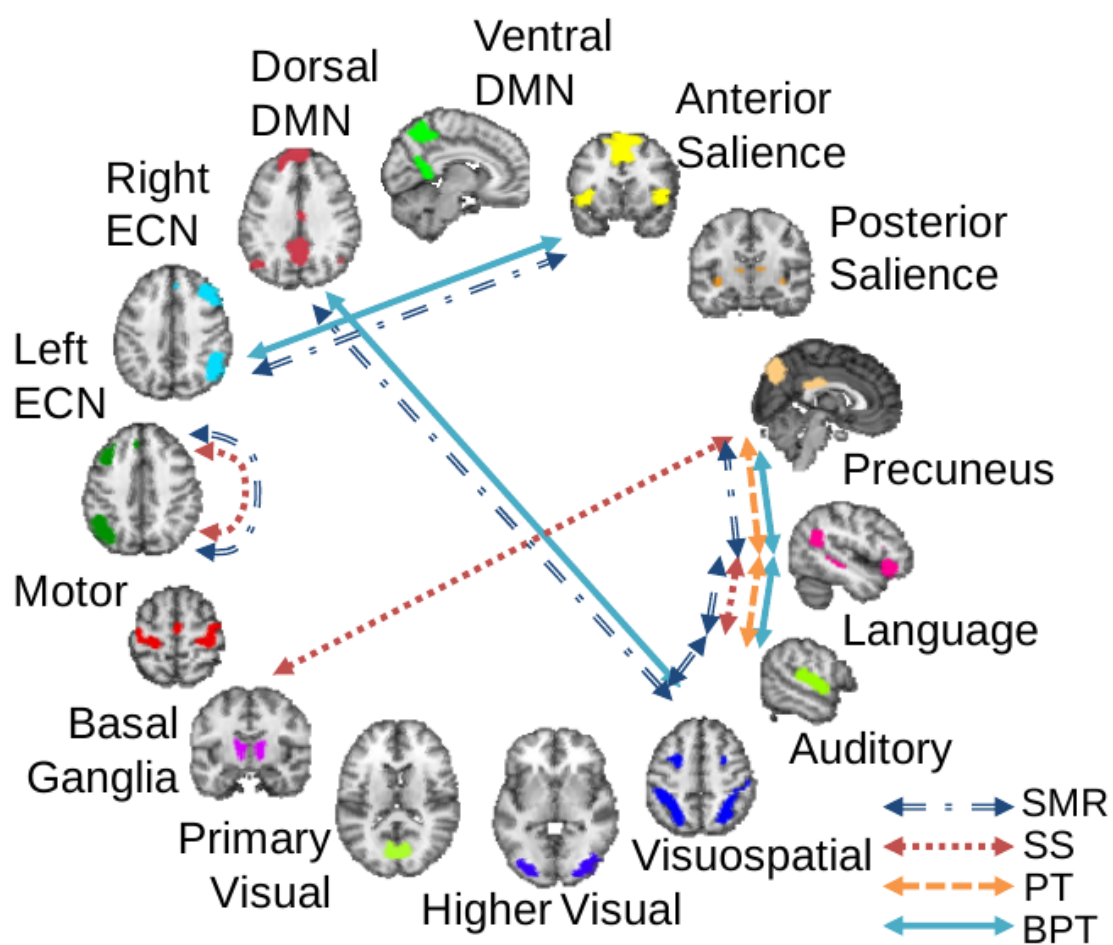


Figure 6.

easily amenable to such an examination: Often there is no apparent structure. Indeed, this is not only a prediction task, but also an inverse problem that calls for adequate regularization. We address this challenge by introducing a convex region-selecting penalty. Our penalty combines total-variation regularization, enforcing spatial contiguity, and l_1 regularization, enforcing sparsity, into one group: Voxels are either active with non-zero spatial derivative or zero with inactive spatial derivative. This leads to segmenting contiguous spatial regions (inside which the signal can vary freely) against a background of zeros. Such segmentation of medical images in a target-informed manner is an important analysis tool. On several prediction problems from brain MRI, the penalty shows good segmentation. Given the size of medical images, computational efficiency is key. Keeping this in mind, we contribute an efficient optimization scheme that brings significant computational gains.

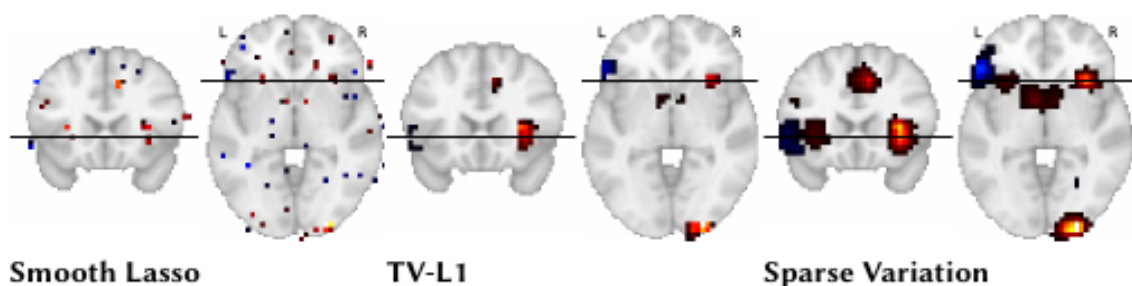


Figure 7.

Weight vectors from estimating gain on the mixed gambles task for three sparse methods: Graphnet, TV-11 and Sparse Variation. This inter- subject analysis shows broader regions of activation. Mean correlation scores on held out data: Graphnet: 0.128, TV-11 : 0.147, Sparse Variation: 0.149. One can see that both TV-11 and Sparse Variation regularizations yield more interpretable patterns than Graphnet.

More information can be found in [40].

7.6. Improving sparse recovery on structured images with bagged clustering

The identification of image regions associated with external variables through discriminative approaches yields ill-posed estimation problems. This estimation challenge can be tackled by imposing sparse solutions. However, the sensitivity of sparse estimators to correlated variables leads to non-reproducible results, and only a subset of the important variables are selected. In this paper, we explore an approach based on bagging clustering-based data compression in order to alleviate the instability of sparse models. Specifically, we design a new framework in which the estimator is built by averaging multiple models estimated after feature clustering, to improve the conditioning of the model. We show that this combination of model averaging with spatially consistent compression can have the virtuous effect of increasing the stability of the weight maps, allowing a better interpretation of the results. Finally, we demonstrate the benefit of our approach on several predictive modeling problems.

Z-score obtained across bootstraps for two discriminative tasks, using the candidate approaches. Higher values hint at lower variability across bootstrap replications. SCLR decreases the variability and yields larger standardized effects.

More information can be found in [42].

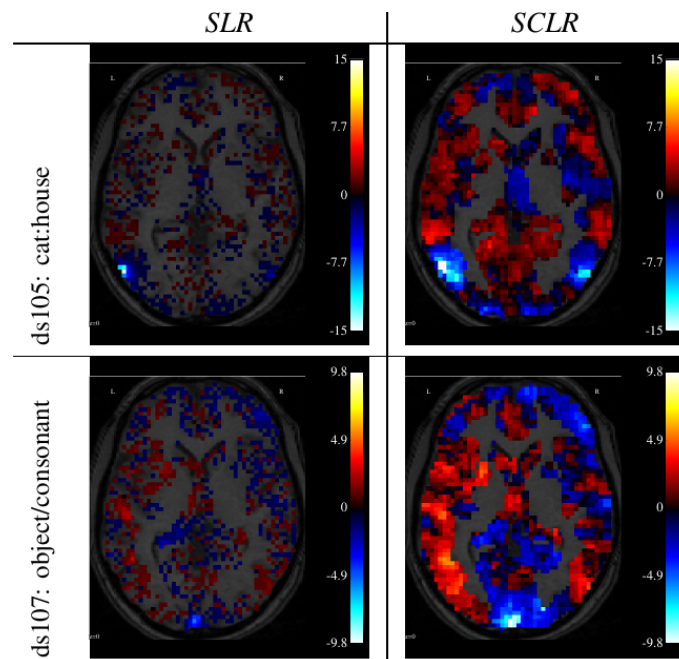


Figure 8.

7.7. Integrating Multimodal Priors in Predictive Models for the Functional Characterization of Alzheimer’s Disease

Functional brain imaging provides key information to characterize neurodegenerative diseases, such as Alzheimer’s disease (AD). Specifically, the metabolic activity measured through fluorodeoxyglucose positron emission tomography (FDG-PET) and the connectivity extracted from resting-state functional magnetic resonance imaging (fMRI), are promising biomarkers that can be used for early assessment and prognosis of the disease and to understand its mechanisms. FDG-PET is the best suited functional marker so far, as it gives a reliable quantitative measure, but is invasive. On the other hand, non-invasive fMRI acquisitions do not provide a straightforward quantification of brain functional activity. To analyze populations solely based on resting-state fMRI, we propose an approach that leverages a metabolic prior learned from FDG-PET. More formally, our classification framework embeds population priors learned from another modality at the voxel-level, which can be seen as a regularization term in the analysis. Experimental results show that our PET-informed approach increases classification accuracy compared to pure fMRI approaches and highlights regions known to be impacted by the disease.

Overview of the proposed classification pipeline: The inputs are ROI-to-voxel connectivities computed from the rs-fMRI time-series. FDG-PET model weights are integrated as prior for the classification. Then, predictions of all ROIs are the inputs of a stacking model to predict the clinical group.

More information can be found in [44].

7.8. Inverse problems with time-frequency dictionaries and Gaussian non-white noise

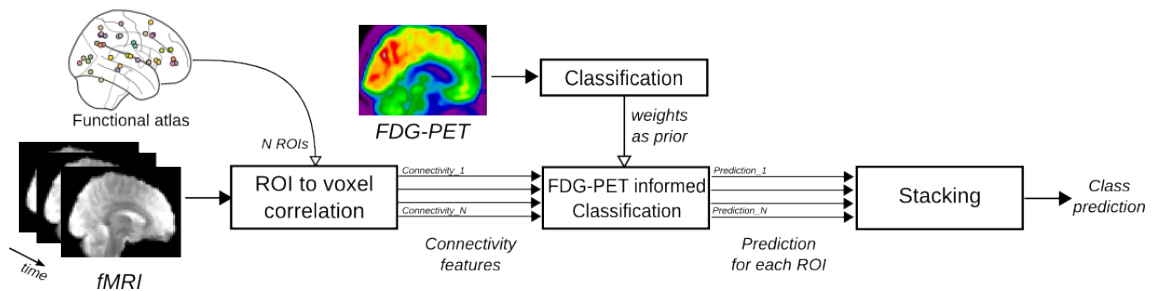


Figure 9.

Sparse regressions to solve ill-posed inverse problems have been massively investigated over the last decade. Yet, when noise is present in the model, it is almost exclusively considered as Gaussian and white. While this assumption can hold in practice it rarely holds when observations are time series as they are corrupted by auto-correlated and colored noise. In this work we study sparse regression under the assumption of non white Gaussian noise and explain how to run the inference using proximal gradient methods. We investigate an application in brain imaging: the problem of source localization using magneto- and electroencephalography (M/EEG) which allow functional brain imaging with high temporal resolution. We use a time-frequency representation of the source waveforms and a sparse regularization which promotes focal sources with smooth and transient activations. Our approach is evaluated using simulations comparing it to strategies that assume the noise is white or to simple prewhitening.

More information can be found in [30].

7.9. Variable density sampling based on physically plausible gradient waveform. Application to 3D MRI angiography

Performing k-space variable density sampling is a popular way of reducing scanning time in Magnetic Resonance Imaging (MRI). Unfortunately, given a sampling trajectory, it is not clear how to traverse it using gradient waveforms. In this paper, we actually show that existing methods can yield large traversal time if the trajectory contains high curvature areas. Therefore, we consider here a new method for gradient waveform design which is based on the projection of unrealistic initial trajectory onto the set of hardware constraints. Next, we show on realistic simulations that this algorithm allows implementing variable density trajectories resulting from the piecewise linear solution of the Travelling Salesman Problem in a reasonable time. Finally, we demonstrate the application of this approach to 2D MRI reconstruction and 3D angiography in the mouse brain.

Full k-space acquisition with an EPI sequence (a) and corresponding reference image (f). Comparison between an exact parameterization of the TSP trajectory (b) and projection from Travelling Salesman Problem trajectory onto the set of constraints (c),(d). In experiments (b,c), the number of measured locations is fixed to 9% ($r = 11.2$), whereas in (b,d), the time to traverse the curve is fixed to 62 ms. (e): Spiral trajectory with acquisition of the k-space center. (g-j): Reconstructed images corresponding to sampling strategies (b-e).

More information can be found in [38].

7.10. A projection method on measures sets.

We consider the problem of projecting a probability measure π on a set MN of Radon measures. The projection is defined as a solution of the following variational problem:

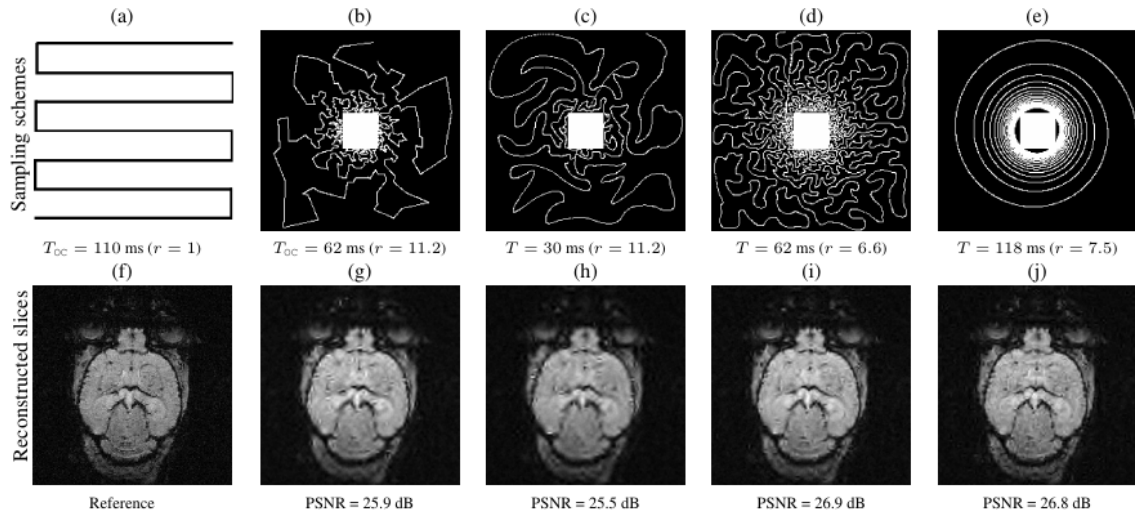


Figure 10.

$$\inf_{\mu \in \mathcal{M}_N} \|h(\mu - \pi)\|_2^2,$$

where $h \in L^2(\Omega)$ is a kernel, $\Omega \subset \mathbb{R}^d$ and denotes the convolution operator. To motivate and illustrate our study, we show that this problem arises naturally in various practical image rendering problems such as stippling (representing an image with N dots) or continuous line drawing (representing an image with a continuous line). We provide a necessary and sufficient condition on the sequence $(\mathcal{M}_N)_{N \in \mathbb{N}}$ that ensures weak convergence of the projections $(\mu_N^*)_{N \in \mathbb{N}}$ to π . We then provide a numerical algorithm to solve a discretized version of the problem and show several illustrations related to computer-assisted synthesis of artistic paintings/drawings.

Projection of a lion image onto $P_N^{1,\infty}$ with $N = 8,000$. The figure depicts the resulting line with several values of the iterates of our Algorithm.

More information can be found in [55].

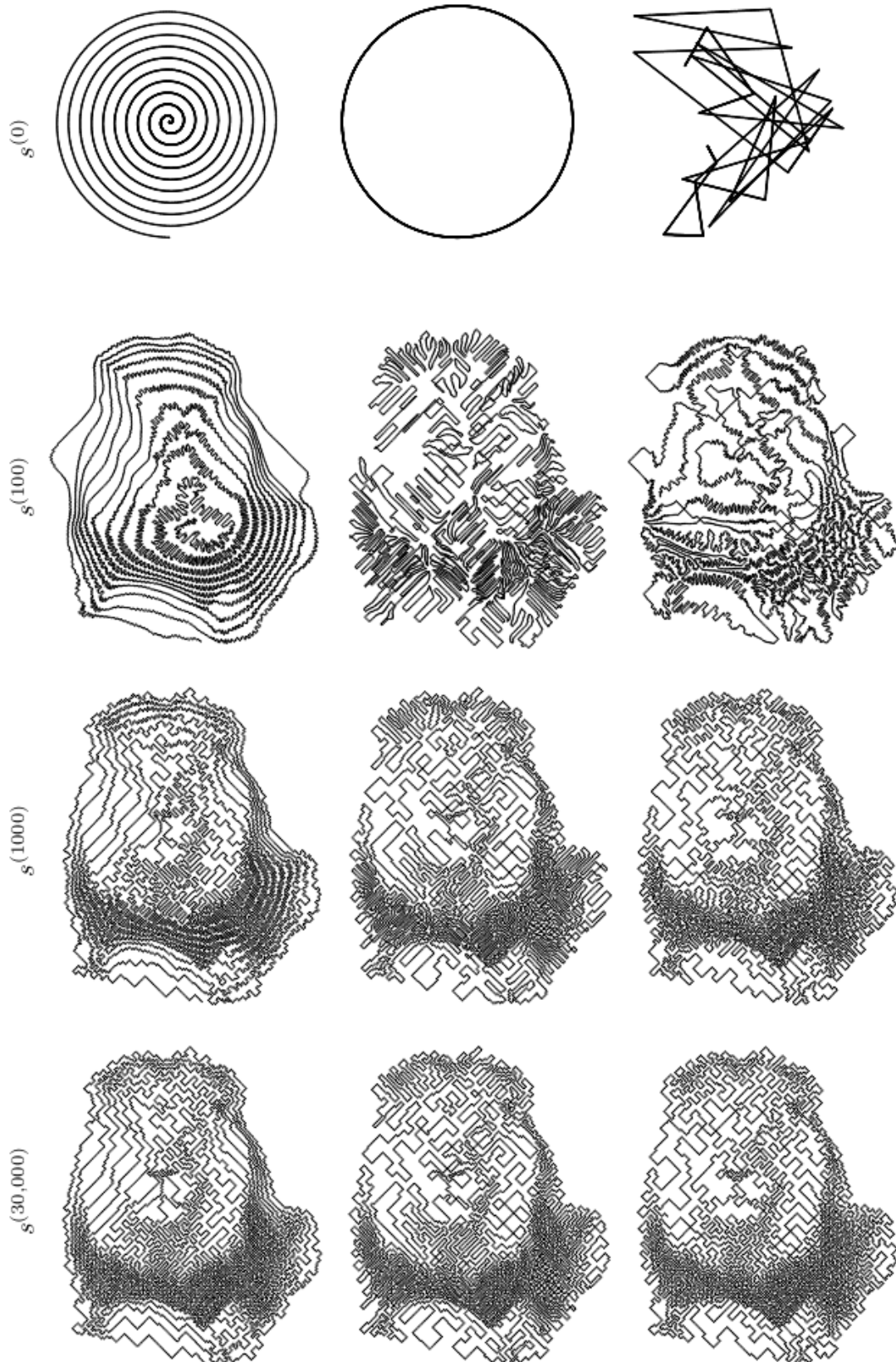


Figure 11.

PLEIADE Team

6. New Results

6.1. Inference of metabolic networks

Participants: David Sherman [correspondant], Razanne Issa, Pascal Durrens.

We are particularly interested in incremental modeling of metabolic networks, where the target organism to be modeled is demonstrably similar to other organisms for which whole or partial models are available. The other organisms are typically strains of the same species as the target, or species with a close phylogenetic relation to the target species. The similarity is measured genomically at different scales: sequence polymorphisms, expansions and contractions in conserved protein families, and genome rearrangements. We have defined and refined two complementary methods for inferring metabolic models for target species.

In the same way that comparative analysis of genomes and proteomes makes it possible to define protein families that summarize protein-coding genes into phyletic patterns [24], comparative analysis of related metabolic models makes it possible to define network generalizations [26] that factor families of reactions and metabolites into summary graphs that preserve stoichiometry. These summaries can be used for expert curation and visualization [5]. An online demonstration tool is made available at <http://mimoza.bordeaux.inria.fr/>.

Starting from an existing reference metabolic network and measures of similarity between the reference and the target organisms' genomes, we can use knowledge-based inference to rewrite the reference network based on these differences, and thus obtain a draft network for the metabolisms of the target organism [2]. This rewriting, formalized in the Pantograph system, can be extended to an abductive logic framework as described in Razanne Issa's thesis [19]. Current work aims at extending the Pantograph and ab-Pantograph frameworks to leverage reaction classifications obtained by network generalization.

6.2. Bio-medicine and biotechnology

Participants: Pascal Durrens [correspondant], David Sherman.

6.2.1. Genome assembly for bio-medicine

We performed the assembly of the *Clavispora lusitaniae* (aka *Candida lusitaniae*) genome. Yeasts from the genus *Candida* are opportunistic human pathogens in immunocompromised patients, linked to a high mortality rate. Although *Candida albicans* is the major pathogen, related species are more and more isolated, such as *Clavispora lusitaniae* which is responsible for candidaemia in newborn babies and in onco-hematology patients.

Even though the genome of a *Clavispora lusitaniae* strain (ATCC 42720), isolated from a patient, has already been sequenced by the Broad Institute, we achieved the genome assembly of the wild type reference strain (CBS 6936) as patient isolates tend to harbor genome modifications. The assembly was computed from Illumina reads with a coverage of 30X, using the MINIA assembler from Inria GENSCALE team. We also looked for single nucleotide polymorphisms (SNPs) in the reads coming from 3 hypovirulent mutants impaired in the beta oxidation metabolic pathway. Some detected SNPs are now under experimental validation and we are going to make a Genome Announcement for the CBS 6936 genome.

6.2.2. Transcriptome assembly for bio-technology

We carried out the assembly of transcriptomes from different tissues of the African oil palm tree *Elaeis guineensis*. The goal of this project is twofold: (i) Select the most relevant genes involved in oil synthesis in order to implement heterologous expression of some of these genes in a cultivated plant recipient such as tobacco. Preliminary results on heterologous expression of 2-3 key genes/factors ended in 15% of dry weight of oil synthesis. New expression technology allows for simultaneous expression of 15-20 genes. Identifying the best candidates for co-expression will permit efficient heterologous oil synthesis. (ii) Identify the polymorphism of genes in a panel of 25 wild type isolates and of 5 production lineages of *Elaeis guineensis* in relation to the oil yield in different environmental conditions. In addition to a high variability of oil quantity (1-12 tons/ha/year), the relative amount of unsaturated fatty acids spans widely (15-55% dry weight) among the 30 *Elaeis guineensis* strains. Identification of polymorphisms will pave the way to genome-wide association genetics (GWAS) for the improvement of the oil resource.

In a first step, we produced assembled transcriptomes of ca. 300 million reads from 3 tissues (leaf, mesocarp, kernel) coming from a single tree, using state-of-the-art assembler TRINITY. Tuning of the software parameters was performed on the Inria PLAFRIM computation platform. About 20% of the assembled sequences revealed to be tissue-specific. Computation of the protein sequences deduced from the assembled transcripts gave a protein repertoire which was annotated using related sequences available in public databases. These transcript and protein sets will be used as a framework in the polymorphism studies.

6.3. Biodiversity and ecology

Participants: Alain Franc [correspondant], Jean-Marc Frigerio, Philippe Chaumeil, Razanne Issa, Leyla Mirvakhabova.

Our main activity has been code development in the framework of a research project with ONEMA, and preparing future development. Code development has been fostered with the work of Razanne Issa (CDD ONEMA) in the last three months of 2015, and preparation of further development has been fostered by welcoming Leyla Mirvakhabova (L3, National Research University Higher School of Economics (NRU HSE), Moscow, mathematics). Declic is a python library providing tools for analysing molecular data for biodiversity studies. The main object is a distance matrix, from which one can either build a point cloud in a Euclidean space with distances between points as close as possible from distances between reads (Multidimensional scaling), or to build a graph with edges between reads when their distance is smaller than a given threshold. Meanwhile, the team has developed the network around the Galaxy server where an early version of tools has been installed and made available, especially with SLU at Uppsala (Maria Khalert). Alain Franc has developed a collaboration with Olivier Coulaud and Pierre Blanchard (Hiepac) for efficient computation of eigenvectors and eigenvalues of large, dense and symmetric matrices, needed for scaling in Multidimensional scaling.

The work of Razanne Issa has made it possible to extend the declic library in the direction of machine learning, by incorporating tools from support vector machines through library sklearn. This development will be pursued in 2016. The work of Leyla Mirvakhabova has permitted a first incursion into topological data analysis as a possible approach for studying the shape of point clouds produced by multidimensional scaling. The collaboration with NRU HSE on this topic will be pursued in 2016.

POPIX Team

7. New Results

7.1. Model identifiability

We have discussed the question of model identifiability within the context of nonlinear mixed effects models. Although there has been extensive research in the area of fixed effects models, much less attention has been paid to random effects models. In this context we distinguish between theoretical identifiability, in which different parameter values lead to non-identical probability distributions, structural identifiability which concerns the algebraic properties of the structural model, and practical identifiability, whereby the model may be theoretically identifiable but the design of the experiment may make parameter estimation difficult and imprecise. We have explored a number of pharmacokinetic models which are known to be non-identifiable at an individual level but can become identifiable at the population level if a number of specific assumptions on the probabilistic model hold. Essentially if the probabilistic models are different, even though the structural models are non-identifiable, then they will lead to different likelihoods.

7.2. Model of tumor growth

Both molecular profiling of tumors and longitudinal tumor size data modeling are relevant strategies to predict cancer patients' response to treatment. Herein we have proposed a model of tumor growth inhibition integrating a tumor's genetic characteristics that successfully describes the time course of tumor size in patients with low-grade gliomas treated with first-line temozolomide chemotherapy. The model captures potential tumor progression under chemotherapy by accounting for the emergence of tissue resistance to treatment following prolonged exposure to temozolomide. Using information on individual tumors' genetic characteristics, in addition to early tumor size measurements, the model was able to predict the duration and magnitude of response, especially in those patients in whom repeated assessment of tumor response was obtained during the first 3 months of treatment. Combining longitudinal tumor size quantitative modeling with a tumor's genetic characterization appears as a promising strategy to personalize treatments in patients with low-grade gliomas

7.3. Methods for PDEs based model

We have extended the methodologies previously developed for ordinary differential equations (ODE) to partial differential equations (PDE). A finite element method solver for a given family of PDEs has been developed. This solver can now be used with a prototype version of Monolix, a platform for population modeling of longitudinal data. We have implemented the well-known Lagrange finite element method in one, two and three dimensions of the space.

REO Project-Team

7. New Results

7.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Matteo Aletti, Faisal Amlani, Benoit Fabrèges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Mikel Landajuela Larma, Damiano Lombardi, Marina Vidrascu.

In [55] we present a numerical study in which several partitioned solution procedures for incompressible fluid-structure interaction are compared and validated against the results of an experimental FSI benchmark. The numerical methods discussed cover the three main families of coupling schemes: strongly coupled, semi-implicit and loosely coupled. Very good agreement is observed between the numerical and experimental results. The comparisons confirm that strong coupling can be efficiently avoided, via semi-implicit and loosely coupled schemes, without compromising stability and accuracy.

In [14] we introduce a Nitsche-XFEM method for fluid-structure interaction problems involving a thin-walled elastic structure (Lagrangian formalism) immersed in an incompressible viscous fluid (Eulerian formalism). The fluid domain is discretized with an unstructured mesh not fitted to the solid mid-surface mesh. Weak and strong discontinuities across the interface are allowed for the velocity and pressure, respectively. The fluid-solid coupling is enforced consistently using a variant of Nitsche's method with cut-elements. Robustness with respect to arbitrary interface intersections is guaranteed through suitable stabilization. Several coupling schemes with different degrees of fluid-solid time splitting (implicit, semi-implicit and explicit) are investigated. A series of numerical tests in 2D, involving static and moving interfaces, illustrates the performance of the different methods proposed.

In [15] we investigated the autoregulation in the retinal haemodynamics by means of three-dimensional simulations. The autoregulation is a key phenomenon from a physiological standpoint, consisting in the ability of the vasculature to control the flow in different pressure conditions. A simplified fluid-structure interaction method was devised in order to render the vessels wall contraction in a large network, with an affordable computational cost. Several test cases were performed on a patient-specific arteriolar network, whose geometry was reconstructed by using fundus camera images. The tests were in agreement with experimental trends and confirm the ability of the approach to reproduce the phenomena involved.

In [33] we study an unsteady nonlinear fluid-structure interaction problem which is a simplified model to describe blood flow through viscoelastic arteries. We consider a Newtonian incompressible two-dimensional flow described by the Navier-Stokes equations set in an unknown domain depending on the displacement of a structure, which itself satisfies a linear viscoelastic beam equation. The fluid and the structure are fully coupled via interface conditions prescribing the continuity of the velocities at the fluid-structure interface and the action-reaction principle. We prove that strong solutions to this problem are global-in-time. We obtain in particular that contact between the viscoelastic wall and the bottom of the fluid cavity does not occur in finite time. To our knowledge, this is the first occurrence of a no-contact result, but also of existence of strong solutions globally in time, in the frame of interactions between a viscous fluid and a deformable structure.

In [27] and [45] we study the effect of wall bending resistance on the motion of an initially spherical capsule freely suspended in shear flow or in a planar hyperbolic flow. We consider a capsule with a given thickness made of a three-dimensional homogeneous elastic material. A numerical method is used to model the coupling of a boundary integral method for the fluids with a shell finite element method for the capsule envelope. For a given wall material, the capsule deformability strongly decreases when the wall bending resistance increases. In addition, if one expresses the same results as a function of the two-dimensional mechanical properties of the mid-surface, which is how the capsule wall is modeled in the thin-shell model, the capsule deformed shape is identical to the one predicted for a capsule devoid of bending resistance. The bending rigidity is found to have

a negligible influence on the overall deformation of an initially spherical capsule, which therefore depends only on the elastic stretching of the mid-surface. Still, the bending resistance of the wall must be accounted for to model the buckling phenomenon, which is observed locally at low flow strength and persists at steady state. We show that the wrinkle wavelength only depends on the bending number, which compares the relative importance of bending and shearing phenomena, and provide the correlation law. Such results can then be used to infer values of the bending modulus and wall thickness from experiments on spherical capsules in simple shear flow.

In [57] we consider the motion of an elastic structure represented by the nonlinear Saint-Venant Kirchhoff model immersed in a compressible fluid modeled by the compressible Navier-Stokes equations. Existence and uniqueness of a regular solution defined locally in time is proved.

7.2. Numerical methods for biological flows

Participants: Chloé Audebert, Benoit Fabrèges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Sanjay Pant, Marc Thiriet, Irène Vignon-Clementel.

In [37] we present a closed-loop global lumped parameter model for pre stage-II single-ventricle physiology. This model, which is built on a fibre mechanics based description of the heart chambers, benefits from a novel method to describe regurgitant valves. As many as 33 model parameters are estimated from uncertain clinical measurements in two patients—with and without atrioventricular valve regurgitation—through the method of data assimilation. Results are validated qualitatively through measurements and clinical estimates that were not included in the parameter estimation procedure. The methods are shown to successfully capture patient-specific clinical observations such as double peaked nature of valvular flows and abnormalities in electrocardiogram readings.

In [39] we propose a methodology for full propagation of uncertainty from clinical data to model results that enables estimation of the confidence associated with model predictions. We illustrate this problem in a pre stage-II single-ventricle physiology, for which coherence of simulations and clinical data indicated that the flow split to the right lung was highly uncertain. We want to assess here how such uncertainty translates into surgical planning of removing the stenosis or not. Taking into account the effect of the rest of the circulation is also studied in the uncertainty propagation.

In [21] 3D blood flow simulations are carried out for the design of a stented valve reducer in enlarged ventricular outflow tracts. Different device designs are built and compared with the initial device-free state, or with the reducer alone. Results suggest that pressure loss is higher for the reducer alone than for the full device, and that the latter successfully restores hemodynamics to a healthy state. Pressure forces on the reducer and on the valve have the same magnitudes. Migration would occur towards the right ventricle rather than the pulmonary arteries.

In [44] we aim at developing a mathematical model in order to reproduce hemodynamics changes due to liver ablation surgeries. First, a 0D closed-loop model is developed, to simulate hepatectomy and compute post-operative average values. Due to the closed loop, the surgery impact both on and from the whole circulation can be captured, including bleeding and infusion. Then, a one-dimensional artery model is implemented to improve the closed-loop model and simulate better the changes in arterial waveforms due to surgery.

In [54] we investigate the spatial and time discretization of the transient Oseen equations. Finite elements with symmetric stabilization in space are combined with several time-stepping schemes (monolithic and fractional-step). Quasi-optimal (in space) and optimal (in time) error estimates are established for smooth solutions in all flow regimes. We first analyze monolithic time discretizations using the Backward Differentiation Formulas of order 1 and 2 (BDF1 and BDF2). We derive a new estimate on the time-average of the pressure error featuring the same robustness with respect to the Reynolds number as the velocity estimate. Then, we analyze fractional-step pressure-projection methods using BDF1. The stabilization of velocities and pressures can be treated either implicitly or explicitly. Numerical results illustrate the main theoretical findings.

In [26] we study the effects of inserted needle on the subcutaneous interstitial flow. A goal is to describe the physical stress affecting cells during acupuncture treatment. The model consists of the convective Brinkman equations to describe the flow through a fibrous medium. Numerical studies in FreeFem++ are performed to illustrate the acute physical stress developed by the implantation of a needle that triggers the physiological reactions of acupuncture. We emphasize the importance of numerical experiments for advancing in modeling in acupuncture. In [40] we show that the acupoint must contain a highly concentrated population of mastocytes (e.g., very-high-amplitude, small-width Gaussian distribution) to get an initial proper response. Permanent signaling is provided by chemotaxis and continuous recruitment of mastocytes. Therefore, the density and distribution of mastocytes are crucial factors for efficient acupuncture as well as availability of circulating and neighboring pools of mastocytes.

In [61] we carry out a three-dimensional blood flow simulation through a complete macrovascular circuit, the cerebral venous network, rather than using reduced order simulation and partial vascular network. The bio-mechanical modeling step is carefully performed and leads to the description of the flow governed by the Navier-Stokes equations for an incompressible viscous fluid. We then numerically solve the equations with a free finite element software in five meshes of a realistic geometry obtained from medical images to prove the feasibility of the pipeline. Some particularities of the venous network, as asymmetry for example, are discussed.

7.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Damiano Lombardi.

In [58] we investigate the monodomain equation which describes the evolution of the cardiac electrical potential and which corresponds to a coupled system involving a reaction-diffusion equation and an ordinary differential equation. Lipschitz stability inequalities are shown for the identification of some parameters of the model from measurements on the cardiac potential and the ionic variable.

In [32] we studied the application of a Reduced-Order Modeling method (Approximated Lax Pairs) to the solution of the partial differential equations describing the polarisation of tissues. Due to the complexity of the scenarios involved and the presence of propagating waves, the performances of the standard methods proposed in the literature to provide a low computational cost solution are not always satisfactory. The ALP method consists of the construction of an adaptive time dependent basis that diagonalises, at each time, a Schrödinger-type operator. Its application to several 2D and 3D test-cases on the equations arising in electrophysiology was investigated, showing that the performances of the method in terms of speed-up and accuracy are promising.

In [62] we considered the simulation of full cycles of the electrical activity of the heart and the corresponding body surface potential. The model is based on a realistic torso and heart anatomy, including ventricles and atria. One of the specificities of our approach is to model the atria as a surface, which is the kind of data typically provided by medical imaging for thin volumes. The bidomain equations are considered in their usual formulation in the ventricles, and in a surface formulation on the atria. Two ionic models are used: the Courtemanche-Ramirez-Nattel model on the atria, and the " Minimal model for human Ventricular action potentials " (MV) by Bueno-Orovio, Cherry and Fenton in the ventricles. The heart is weakly coupled to the torso by a Robin boundary condition based on a resistor-capacitor transmission condition. Various ECGs are simulated in healthy and pathological conditions (left and right bundle branch blocks, Bachmann's bundle block, Wolff-Parkinson-White syndrome). To assess the numerical ECGs, we use several qualitative and quantitative criteria found in the medical literature. Our simulator can also be used to generate the signals measured by a vest of electrodes. This capability is illustrated at the end of the article.

In [24] we address the inverse problem of electrocardiography from a new perspective, by combining electrical and mechanical measurements. Our strategy relies on the definition of a model of the electromechanical contraction which is registered on ECG data but also on measured mechanical displacements of the heart tissue typically extracted from medical images. In this respect, we establish in this work the convergence of a sequential estimator which combines for such coupled problems various state of the art sequential data

assimilation methods in a unified consistent and efficient framework. Indeed we aggregate a Luenberger observer for the mechanical state and a Reduced Order Unscented Kalman Filter applied on the parameters to be identified and a POD projection of the electrical state. Then using synthetic data we show the benefits of our approach for the estimation of the electrical state of the ventricles along the heart beat compared with more classical strategies which only consider an electrophysiological model with ECG measurements. Our numerical results actually show that the mechanical measurements improve the identifiability of the electrical problem allowing to reconstruct the electrical state of the coupled system more precisely. Therefore, this work is intended to be a first proof of concept, with theoretical justifications and numerical investigations, of the advantage of using available multi-modal observations for the estimation and identification of an electromechanical model of the heart.

7.4. Lung and respiration modeling

Participants: Laurent Boudin, Muriel Boulakia, Céline Grandmont, Jessica Oakes, Nicolas Pozin, Irène Vignon-Clementel.

In silico models of flow and transport in the lung are increasingly being used to predict regional deposition in healthy and diseased lungs. However, very few models have been validated with in vivo human or animal experimental data. In [36], we create a physiologically-based simulation of airflow and particle transport in healthy and emphysematous rat lungs. Excellent agreement between the numerical predictions and experimental data is found for the healthy lungs. However, the numerical predictions are unable to predict the experimental findings of enhanced deposition in the normal regions of the emphysematous lungs and thus more sophisticated models of transport in the deep regions of the lung are needed. This is what is being explored in [42], where interactions of flow and transport between 3D upper-parts and 1D downstream respiratory trees are captured for inspiration and expiration for the first time.

While several groups have investigated detailed flow and particle transport in the acinar regions of the healthy lung, little is currently known about diseased acini. In [35] we perform numerical simulations of flow and transport in healthy and emphysematous acini. As the alveolar septa is deteriorated in emphysema there is less surface area available for particles to deposit on. Therefore, fewer particles deposit in the diseased models. In addition, we find that particle deposition is more heterogeneously distributed in emphysema, a phenomenon that was also found in the in vivo animal experiments.

7.5. Methods for the interaction data - simulation

Participants: Jean-Frédéric Gerbeau, Damiano Lombardi, Sanjay Pant, Irène Vignon-Clementel.

In [38] we proposed an information theoretical framework to study the practical identifiability of dynamical systems. The fundamental question arising in parameter estimation problems is whether, given a set of observations of the system, it is possible to retrieve the parameters values. The method proposed exploits a database of direct numerical simulations and study the parameters-to-observables map by means of differential entropies. Contrary to other approaches proposed in the literature it is not restricted to ordinary differential equations and take the experimental noise into account. Several test cases were performed on a large spectrum of bio-physical systems, providing promising results.

In [60] we studied a differential entropy estimator based on kp -neighbours, aiming at applying a Bayesian framework and some information-theoretic ideas to inverse problems. The goal of this work is to estimate the Shannon differential entropy in high dimensional settings, in possible presence of functional or nearly functional dependences. A modification of the Kozachenko-Leonenko estimator is proposed, consisting of introducing a local gaussian approximation to the probability measure. Test-cases were performed to assess the properties of the method and to compare its performances with other methods proposed in the literature.

The articles [37], presented in the section about biological flows, and [24], presented in the section about electrophysiology, also present methods concerning the interaction data - simulation.

7.6. Miscellaneous

Participants: Laurent Boudin, Irène Vignon-Clementel.

In [34] we develop a quantitative single cell-based model for multi-cellular tumor spheroids of a specific lung cancer cell line, growing under various nutrient conditions: we confront the simulations performed with this model with data on the growth kinetics and spatial labeling patterns for cell proliferation, extracellular matrix, cell distribution and cell death. We stepwise arrive at a model that mimics the spheroid growth under two conditions, and can predict two other ones. The number of mechanisms the model contains is necessary and sufficient to explain the data.

In [19] we consider a kinetic model describing some mechanisms of opinion formation in the framework of referendums, where the individuals, who can interact between themselves and modify their opinion by means of spontaneous self-thinking, are moreover under the influence of mass media. We study, at the numerical level, both the transient and the asymptotic regimes. In particular, we point out that a plurality of media, with different orientations, is a key ingredient to allow pluralism and prevent consensus. The forecasts of the model are compared to some surveys related to the Scottish independence referendum of 2014.

In [56] we review various results on the compactness of the linearized Boltzmann collision operator and of its generalization to mixtures of non-reactive monatomic gases.

SAGE Project-Team

7. New Results

7.1. Numerical algorithms

7.1.1. Introduction to computational linear algebra

Participant: Jocelyne Erhel.

Publications: [23]

Abstract: The book "Introduction to Computational Linear Algebra" presents classroom-tested material on computational linear algebra and its application to numerical solutions of partial and ordinary differential equations. The book is designed for senior undergraduate students in mathematics and engineering as well as first-year graduate students in engineering and computational science.

The text first introduces BLAS operations of types 1, 2, and 3 adapted to a scientific computer environment, specifically MATLAB. It next covers the basic mathematical tools needed in numerical linear algebra and discusses classical material on Gauss decompositions as well as LU and Cholesky's factorizations of matrices. The text then shows how to solve linear least squares problems, provides a detailed numerical treatment of the algebraic eigenvalue problem, and discusses (indirect) iterative methods to solve a system of linear equations. The final chapter illustrates how to solve discretized sparse systems of linear equations. Each chapter ends with exercises and computer projects.

7.1.2. Hybrid algebraic sparse linear solvers

Participants: Jocelyne Erhel, David Imberti.

Grants and projects: EXA2CT 9.2.1 , EoCoE 9.2.2 , C2S@EXA 9.1.2

Publications: in preparation.

Abstract: Sparse linear systems arise in computational science and engineering. The goal is to reduce the memory requirements and the computational cost, by means of high performance computing algorithms. Krylov methods combined with Domain Decomposition are very efficient for both fast convergence and fast computations.

7.1.3. Hastings-Metropolis Algorithm on Markov Chains for Small-Probability Estimation

Participant: Lionel Lenôtre.

Grants: H2MNO4 9.1.1

Publications: [12]

Abstract: Shielding studies in neutron transport, with Monte Carlo codes, yield challenging problems of small-probability estimation. The particularity of these studies is that the small probability to estimate is formulated in terms of the distribution of a Markov chain, instead of that of a random vector in more classical cases. Thus, it is not straightforward to adapt classical statistical methods, for estimating small probabilities involving random vectors, to these neutron-transport problems. A recent interacting-particle method for small-probability estimation, relying on the Hastings-Metropolis algorithm, is presented. It is shown how to adapt the Hastings-Metropolis algorithm when dealing with Markov chains. A convergence result is also shown. Then, the practical implementation of the resulting method for small-probability estimation is treated in details, for a Monte Carlo shielding study. Finally, it is shown, for this study, that the proposed interacting-particle method considerably outperforms a simple Monte Carlo method, when the probability to estimate is small.

7.1.4. A Strategy for the Parallel Implementations of Stochastic Lagrangian Methods

Participant: Lionel Lenôtre.

Grants and projects: H2MNO4 [9.1.1](#)

Software: PALMTREE [6.5](#)

Publications: [[32](#)]

Abstract: We present some investigations on the parallelization of a stochastic Lagrangian simulation. For the self sufficiency of this work, we start by recalling the stochastic methods used to solve Parabolic Partial Differential Equations with a few physical remarks. Then, we exhibit different object-oriented ideas for such methods. In order to clearly illustrate these ideas, we give an overview of the library PALMTREE that we developed. After these considerations, we discuss the importance of the management of random numbers and argue for the choice of a particular strategy. To support our point, we show some numerical experiments of this approach, and display a speedup curve of PALMTREE. Then, we discuss the problem in managing the parallelization scheme. Finally, we analyze the parallelization of hybrid simulation for a system of Partial Differential Equations. We use some works done in hydrogeology to demonstrate the power of such a concept to avoid numerical diffusion in the solution of Fokker-Planck Equations and investigate the problem of parallelizing scheme under the constraint entailed by domain decomposition. We conclude with a presentation of the latest design that was created for PALMTREE and give a sketch of the possible work to get a powerful parallelized scheme.

7.1.5. *About a generation of a log-normal correlated field*

Participants: Jocelyne Erhel, Géraldine Pichot.

Grants: HYDRINV [9.3.3](#) , H2MN04 [9.1.1](#)

Software: GENFIELD [6.1](#)

Publications: [[18](#)].

Abstract: Uncertainty quantification often requires the generation of large realizations of stationary Gaussian random field over a regular grid.

We compare the classical methods used to simulate the field defined by its covariance function, namely the Discrete Spectral method, the Circulant Embedding approach, and the Discrete Karhunen-Loève approximation. We design and implement a parallel algorithm related to the Discrete Spectral method.

7.2. Numerical models and simulations applied to heat transfer

7.2.1. *Small scale modeling of porous media*

Participants: Édouard Canot, Salwa Mansour.

Grants: ECOS Sud Chili (ARPHYMAT project) [9.3.2](#)

Software: GLiMuH [6.2](#)

Publications: [[13](#)]

Conferences: [[20](#)]

Abstract: This study is devoted to the heat transfer between two spherical grains separated by a small gap; dry air is located around the grains and a liquid water meniscus is supposed to be present between them. This problem can be seen as a micro-scale cell of an assembly of solid grains, for which we are looking for the effective thermal conductivity. For a fixed contact angle and according to the volume of the liquid meniscus, two different shapes are possible for the meniscus, giving a “contacting” state (when the liquid makes a true bridge between the two spheres) and a “non-contacting” one (when the liquid is split in two different drops, separated by a thin air layer); the transition between these two states occurs at different times when increasing or decreasing the liquid volume, thus leading to a hysteresis behavior when computing the thermal flux across the domain. We consider also another process where humidity varies, for example during an evaporation or condensation process; in this situation, the shape of the menisci changes a lot, because some liquid bridges may break, and this can strongly affect the effective thermal conductivity. Then, the reorganization of the liquid menisci is predicted, especially their surface area variation; it is an important parameter for a global model of the evaporation phenomenon in wet porous media.

7.2.2. *Inverse problem for determining the thermo-physical properties of a porous media*

Participants: Édouard Canot, Salwa Mansour.

Grants: HYDRINV 9.3.3

Software: TPIP (6.7)

Publications: [15], [27]

Conferences: [22]

Abstract: This study concerns the inverse problem which consists of the estimation of thermophysical properties of the soil knowing the temperature at few selected points of the domain. In order to solve this inverse problem, we used the least square criterion where we try to minimize the error function between real measures and simulated ones. The coupled system composed of the energy equation together with the three sensitivity boundary initial problems resulting from differentiating the basic energy equation with respect to the soil properties must be solved. To overcome the stiffness of our problem (due to the use of Apparent Heat Capacity method), the high nonlinearity of the coupled system and the problem of large residuals we used the Damped Gauss Newton and Levenberg-Marquardt methods. To take into account uncertainties of the position of the sensors, some constraints have been added to the least square problem. Results are good when the number of sensors is sufficiently large.

7.2.3. *Evaporation/Condensation in a wet granular medium: the EWGM model*

Participants: Édouard Canot, Salwa Mansour.

Grants: ECOS Sud Chili (ARPHYMAT project) 9.3.2

Software: HeMaTis (6.4)

Publications: [26], [25]

Abstract: The physical model of the HeMaTiS code (6.4) has been completed by a new variant dedicated to the unsaturated case. The pendular regime concerns the special case where a very few quantity of liquid water is contained in a granular medium. The new model involves seven variables and can be considered as a two-phase two-component one; it contains both air and water, this latter component being liquid or gas. Generally, the diffusive transport of humidity in soils is extremely slow, we numerically show that humidity is convected quickly when the medium is subjected to a strong temperature gradient. The key feature of the thermal process is the simultaneous evaporation and condensation of water near a discontinuity of the liquid layout.

7.3. Models and simulations for skew diffusion

7.3.1. *Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests*

Participant: Géraldine Pichot.

Grants: H2MN04 9.1.1

Software: SBM 6.6

Publications: submitted.

Abstract: We present several benchmark tests for Monte Carlo methods for simulating diffusion in one-dimensional discontinuous media, such as the ones arising the geophysics and many other domains. These benchmarks tests are developed according to their physical, statistical, analytic and numerical relevance. We then perform a systematic study on four numerical methods.

7.3.2. *One-dimensional skew diffusions: explicit expressions of densities and resolvent kernel*

Participants: Lionel Lenôtre, Géraldine Pichot.

Grants: H2MN04 9.1.1

Publications: [31]

Abstract: The study of skew diffusion is of primary concern for their implication in the modeling and simulation of diffusion phenomena in media with interfaces. First, we provide results on one-dimensional processes with discontinuous coefficients and their connections with the Feller theory of generators as well as the one of stochastic differential equations involving local time. Second, in view of developing new simulation techniques, we give a method to compute the density and the resolvent kernel of skew diffusions. Explicit closed-form are given for some particular cases.

7.3.3. Algorithms for the simulation of Feller processes

Participant: Lionel Lenôtre.

Grants and projects: H2MNO4 9.1.1 .

Publications: [34].

Abstract: Two new numerical schemes are created for Skew Diffusions processes. Both algorithms rely on a more generic numerical scheme that can be used for any kind of Feller processes. The proof of convergence for this generic numerical scheme is performed.

7.3.4. Theoretical results on multidimensional Skew Diffusions

Participant: Lionel Lenôtre.

Grants and projects: H2MNO4 9.1.1 .

Publications: [33].

Abstract: Some significant results on the distribution of the marginal processes of multidimensional Skew Diffusions are found together with new formula. In addition, totally analytical proofs of some results and algorithms given by A. Lejay are given.

7.4. Models and simulations for flow and transport in porous fractured media

7.4.1. An adaptive sparse grid method for elliptic PDEs with stochastic coefficients

Participant: Jocelyne Erhel.

Grants and projects: HYDRINV 9.3.3 , H2MN04 9.1.1

Publications: [14].

Abstract: The stochastic collocation method based on the anisotropic sparse grid has become a significant tool to solve partial differential equations with stochastic inputs. The aim is to seek a vector of weights and a convenient level of interpolation for the method. The classical approach uses an a posteriori approach on the solution, which causes an additional prohibitive cost.

In this work, we discuss an adaptive approach of this method to calculate the statistics of the solution. It is based on an adaptive approximation of the *inverse* diffusion parameter. We construct an efficient error indicator which is an upper bound of the error on the solution. In the case of unbounded variables, we use an appropriate error estimation to compute suitable weights for the method. Numerical examples are presented to confirm the efficiency of the approach, and to show that the cost is considerably reduced without loss of accuracy.

7.4.2. A global reactive transport model applied to the MoMaS benchmark

Participant: Jocelyne Erhel.

Grants and projects: H2MN04 9.1.1

Software: GRT3D 6.3

Publications: [19].

Abstract: Reactive transport models are very useful for groundwater studies such as water quality, safety analysis of waste disposal, remediation, and so on. The MoMaS group defined a benchmark with several test cases. We present results obtained with a global method and show through these results the efficiency of our numerical model.

7.4.3. About some numerical models for geochemistry

Participant: Jocelyne Erhel.

Grants and projects: H2MN04 [9.1.1](#)

Publications: [\[16\]](#), [\[17\]](#).

Abstract: Reactive transport models are very useful to study the fate of contaminants in groundwater. These models couple transport equations with geochemistry equations. In this talk, we focus on precipitation and dissolution chemical reactions, because they induce numerical difficulties.

We consider a set of solute species and minerals, with precipitation occurring when a saturation threshold is reached. A challenge is to detect which minerals are dissolved and which minerals are precipitated. This depends on the total quantities of chemical species. We propose an analytical approach to build a phase diagram, which provides the interfaces between the different possible cases. We illustrate our method with three examples arising from brine media and acid mine drainage.

7.4.4. Power-averaging method to characterize and upscale permeability in DFNs

Participants: Jean-Raynald de Dreuzy, Géraldine Pichot.

Publications: [\[21\]](#).

Abstract: In a lot of geological environments, permeability is dominated by the existence of fractures and by their degree of interconnections. Flow properties depend mainly on the statistical properties of the fracture population (length, apertures, orientation), on the network topology, as well as on some detailed properties within fracture planes. Based on an extensive analysis of 2D and 3D DFNs as well as on reference connectivity structures, we investigate the relation between the local fracture structures and the effective permeability. Defined as the relative weight between the two extreme harmonic and arithmetic means, the power-law averaging exponent gives a compact way to compare fracture network hydraulics. It may further lead to some comprehensive upscaling rules.

SERENA Team

7. New Results

7.1. Guaranteed bounds for Laplace eigenvalues and eigenvectors

In [22], we have derived a posteriori error estimates for the Laplace eigenvalue problem. Guaranteed, fully computable, and optimally convergent upper and lower bounds for the first eigenvalue are given. They are valid under explicit, a posteriori conditions on the computational mesh and on the approximate solution. Guaranteed, fully computable, and polynomial-degree robust bounds for the energy error in the approximation of the first eigenvector are derived as well, under the same conditions. Remarkably, all the constants in our theory can be fully estimated, and no convexity/regularity assumption on the computational domain/exact eigenvector(s) is needed. This general result can still be improved when an elliptic regularity assumption is satisfied (with known constants), typically for convex two-dimensional domains. The application of our framework to conforming finite element approximations of arbitrary polynomial degree is provided, along with a numerical illustration on a set of test problems.

SERPICO Project-Team

7. New Results

7.1. Statistical aggregation methods for image denoising and estimation

Participants: Charles Kervrann, Frédéric Lavancier.

We have already proposed a general statistical aggregation method which combines image patches denoised with several commonly-used algorithms [10]. We showed that weakly denoised versions of the input image obtained with standard methods, can serve to compute an efficient patch-based aggregated estimator. In our approach, the Stein's Unbiased Risk Estimator (SURE) is used to evaluate each denoised candidate image patch and to compute the exponential weighted aggregation (EWA) estimator. This year, we adapted this framework (PEWA) to denoise images corrupted by mixed Gaussian-Poisson in 2D fluorescence image sequences.

In this range of work, we have also introduced in [24] a general method to combine estimators in order to produce a better estimate. From a theoretical point of view, we proved that this method is optimal in some sense. It is illustrated on standard statistical problems in parametric and semi-parametric models where the averaging estimator outperforms the initial estimators in most cases. As part of an on-going work, we are applying this method to improve patch-based image denoising algorithms.

Reference: [24]

Collaborators: Paul Rochet (Laboratoire de Mathématiques Jean Leray (LMJL), University of Nantes).

7.2. Image deconvolution algorithms for tagged-RNA and gene localization in live yeast

Participant: Charles Kervrann.

In fluorescence microscopy, the image quality is limited by out-of-focus blur and high noise. Traditionally, image deconvolution is needed to estimate a good quality version of the observed image. The result of deconvolution depends heavily on the choice of the regularization term. The regularization functional should be designed to remove noise while retaining the image structure. In this study, we investigated non quadratic regularization terms to preserve fine details of underlying structures and we studied appropriate optimization algorithms. The deconvolution method has been especially dedicated for 3D high-precision gene localization in cell nuclei [47]. For illustration, tagged gene (green marker) and tagged nucleoporins/nuclear periphery (red marker) are shown in Fig. 3. A noisy and blurred image can affect the nuclear membrane estimation and gene detection and, consequently, the computed related distances.

Collaborators: Giovanni Petrazzuoli (Inserm U944, CNRS UMR 9212, Hôpital Saint-Louis, Paris),
Catherine Dargemont (Inserm U944, CNRS UMR 9212, Hôpital Saint-Louis, Paris),
Jean Salamero (UMR 144 CNRS-Institut Curie, PICT-IBiSA).

7.3. Estimation of the reference point giving the most uniform angular distribution

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

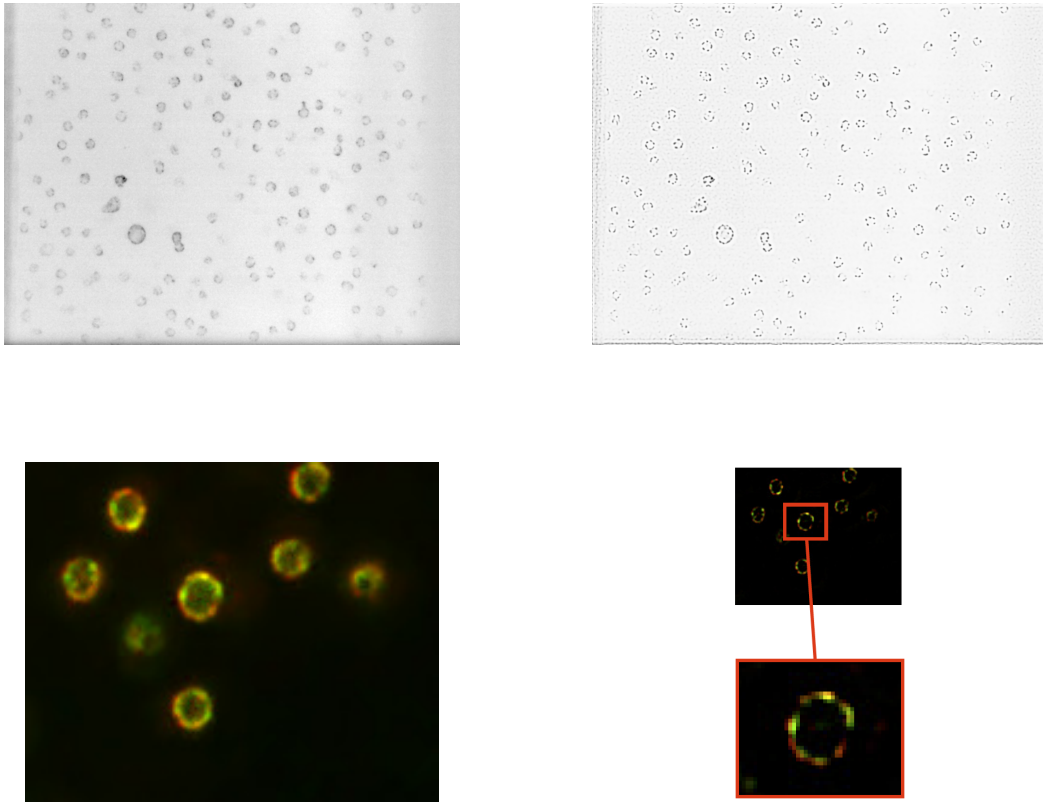


Figure 3. Deconvolution of 3D image depicting tagged gene and tagged nucleoporins/nuclear periphery. First row: deconvolution (right) of a tagged nucleoporin image (left). Second row: blurred image of tagged gene and nucleoporins (left) and zoom-in view of the deblurred image.

Rab11 proteins are trafficking from the Endosomal Recycling Compartment (ERC) to locations in the cell membrane where they eventually fuse. In this study, we assume that the Rab11 positive membranes are uniformly distributed around the ERC at the cell membrane. To test this hypothesis, we estimate the angular distribution of Rab11 positive membranes from several image sequences acquired with a TIRF microscope at the cell membrane level by considering all the points located in the cell as a reference point. We then compute the entropy of angular distribution for each point and estimate the ERC location as the reference point that gives the maximum entropy for the angular distribution (see Fig. 4). These results are very close to the ERC locations manually annotated by experts.

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, PICT-IBiSA),
Jérôme Boulanger (UMR 144 CNRS-Institut Curie),
Liu Zengzhen (UMR 144 CNRS-Institut Curie).

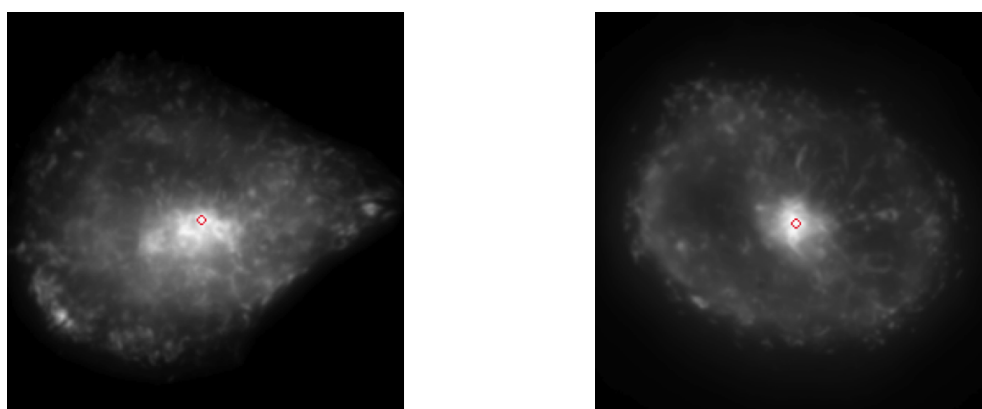


Figure 4. Estimation of reference points (red circles) for Rab11 traffic comparison, registration and quantification.

7.4. Modeling and estimation of protein release and diffusion in TIRFM

Participants: Antoine Basset, Charles Kervrann, Patrick Boutheymy.

We have pursued our work on membrane dynamics, still following a local approach in space and time. We have proposed a new model to account for the full behavior of cargo transmembrane proteins during the vesicle fusion to the plasma membrane at the end of the exocytosis process (see Fig. 5). It combines release and diffusion steps. The former is represented by an exponential decay to account for a continuous release of the proteins from the vesicle to the plasma membrane. We can relax the usual point source assumption, and we name our model the “Small-extent Source with Exponential Decay release” (SSED). An iterative minimization method is used to estimate simultaneously both biophysical parameters, i.e., the release rate and the diffusion coefficient, for every active vesicle detected in the total internal reference fluorescence microscopy (TIRFM) image sequence. Quantitative evaluation has demonstrated the efficiency of the method, which has also allowed us to exhibit differences in the behaviors of Transferrin receptor (TfR) and Langerin proteins.

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, PICT-IBiSA),
Jérôme Boulanger (UMR 144 CNRS-Institut Curie).

7.5. Counting-based particle flux estimation for traffic analysis in live cell imaging

Participants: Thierry Pécot, Charles Kervrann.

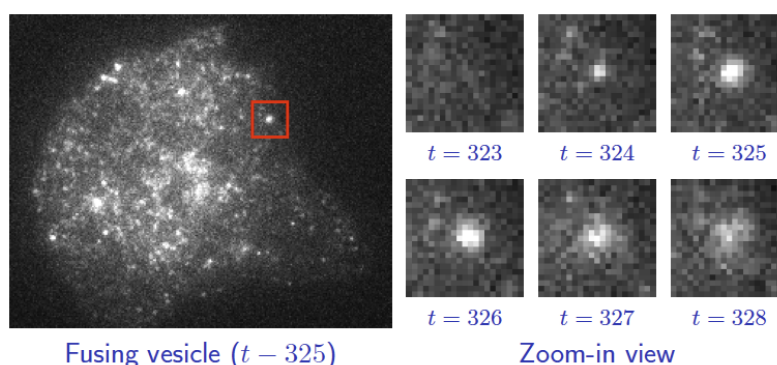


Figure 5. Left: Fusing vesicle (frame in red) in a TIRFM (UMR 144 CNRS-Institut Curie, PICT-IBiSA) sequence (frame 325, 50ms/frame). Right: Zoom-in view of the temporal evolution of the fusing vesicle.

In this study, we have proposed an original traffic analysis approach based on the counting of particles from frame to frame. Object tracking methods or optical flow methods are generally considered to analyze the dynamic contents of intracellular video-microscopy. The suggested method lies between these two well-known approaches. Instead of tracking each moving particle, we estimate fluxes of particles between predefined and adjacent regions. Our three-step counting-based approach is as follows:

- The cell is uniformly partitioned into fixed-size and fixed-shape regions.
- The moving particles are automatically detected using an appropriate algorithm.
- The fluxes are estimated with sparse constraints from an image pair at each time step from temporal variations of the number of particles in each region of the uniform tessellation. Except for some trivial cases, the flux estimation is actually an ill-posed problem and additional constraints are necessary to find the optimal solution.

The problem is formulated as the minimization of a global cost function and the approach allows us to process image sequences with a high number of particles and a high rate of particle appearances and disappearances. We studied the influence of object density, image partition scale, motion amplitude and particle appearances/disappearances in a large variety of simulations. The potential of the method has been demonstrated on real image sequences showing GFP-tagged Rab6 trafficking in confocal microscopy.

Reference: [26]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, PICT-IBiSA),
Jérôme Boulanger (UMR 144 CNRS-Institut Curie).

7.6. Tracking of astral microtubules at the cell cortex

Participants: Frédéric Logé-Munerel, Thierry Pécot, Antoine Basset, Charles Kervrann.

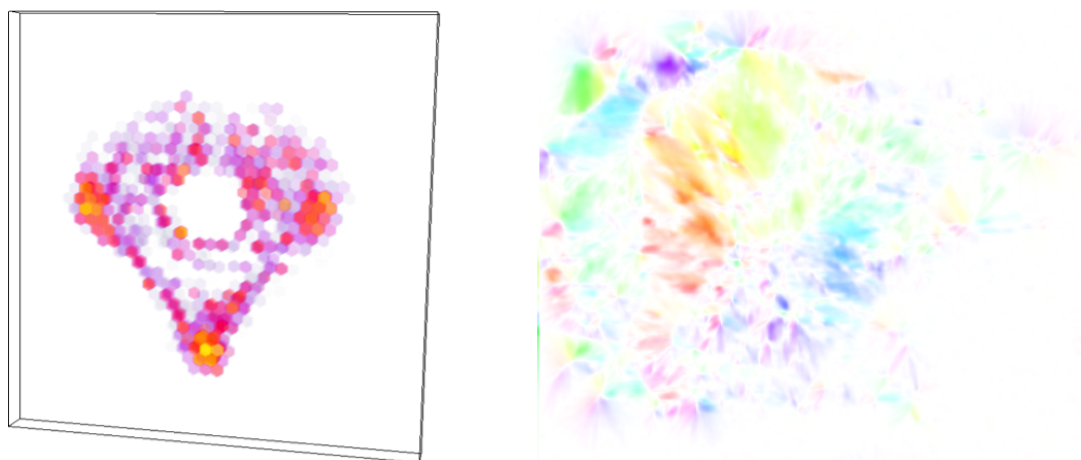


Figure 6. Left: Spatial distribution of GFP-tagged Rab6 vesicle numbers estimated when considering a regular 3D tessellation. Right: Estimation of directional/intensity flows of actin filaments for an image sequence acquired in TIRF microscopy (UMR 144 CNRS-Institut Curie, PICT-IBiSA).

In this study, we are currently interested in the influence of the mechanical properties of astral microtubules in the centering mechanisms of the mitotic spindle, giving it a robust positioning. In their previous studies, the CEDRE group (IGDR Rennes) identified two subpopulations of astral microtubules that either push or pull the cell cortex. To better understand these mechanisms, they acquired image sequences at the cortex level where astral microtubules extremities come to exert forces. In order to characterize the two subpopulations of astral microtubules during the mitosis in the unicellular embryos of *C. Elegans* life span, that is the period during which the microtubule is touching the cell cortex, has to be measured for every single microtubule. A short life span corresponds to a pulling force and a long life span corresponds to a pushing force. Detecting and tracking microtubules at the cell cortex has to be done to collect these measures. This year, F. Logé-Munere (internship Master 1, supervisors: T. Pécot and C. Kervrann) improved the analysis workflow and calibrated the parameters of the algorithms to successfully track the microtubules. This workflow is composed of the ND-SAFIR denoising algorithm [4], the ATLAS detection algorithm [12] and the ASTRE tracking algorithm [56]. The experimental results are currently compared with results obtained by the CEDRE group using the U-track platform [50] (see Fig. 7).

Collaborators: Jacques Pécreaux and H el ene Bouvrais (CEDRE group, IGDR Rennes, CNRS UMR 6290).

7.7. Correlation-based method for membrane diffusion estimation during exocytosis in TIRFM

Participants: Ancageorgiana Caranfil, Antoine Basset, Charles Kervrann.

The dynamics of the plasma membrane of the cell is not fully understood yet; one of the crucial aspects to clarify is the diffusion process during exocytosis. Several observation methods exist, including TIRFM (Total Internal Reflection Fluorescence Microscopy), that has successfully been used to determine the successive steps of exocytosis. However, computing characteristic values for plasma membrane dynamics is problematic, as the experimental conditions have a strong influence on the obtained data and a global model cannot be determined. The goal of this study was to build a correlation-like method to estimate local diffusion parameters in TIRFM images. Using a correlation approach similar to TICS (Temporal Image Correlation Spectroscopy) with an adapted local model, we have developed a novel correlation-based method to estimate the diffusion

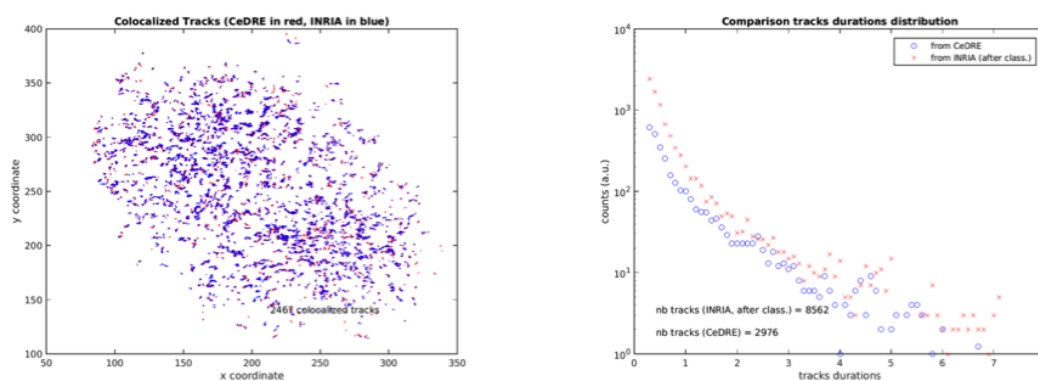


Figure 7.

Microtubule extremities detection and tracking in fluorescence microscopy (embryo of C. Elegans, IGDR - Institute of Genetics and Developmental biology of Rennes, CNRS UMR 6290).

coefficient for every diffusion event in TIRFM images. We turned the non-linear model of the TICS method into a linear one, and made it rely on less parameters than the other estimation methods. Results are excellent for sequences with a good signal-to-noise ratio (see Fig. 8); however, time and space dependencies are introduced with the presence of moderate-to-strong image noise. Although only synthetic images have been used so far, studies of real-life TIRFM images are forthcoming, along with refinements to make the method robust to noise.

Collaborators: Perrine Paul-Gilloteaux and Francois Waharte (UMR 144 CNRS-Institut Curie, PICT-IBiSA).

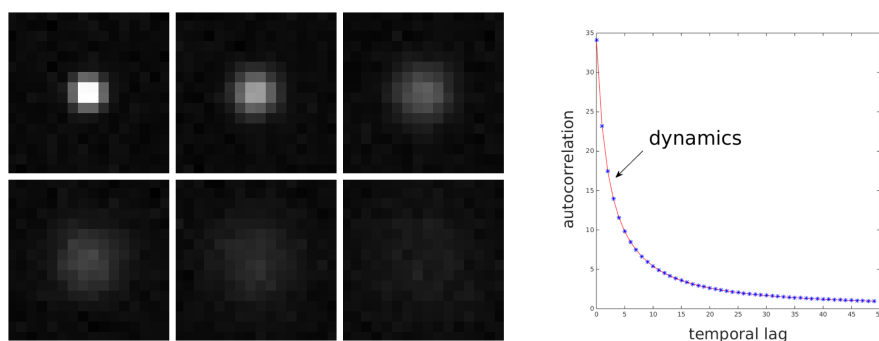


Figure 8. Left: first six instances of a TIRFM image sequence showing a diffusion event. Right: the correlation-based method is applied on the TIRFM sequence; both the computed values of the autocorrelation, for different values of the temporal lag, and the fitting function for these values are represented.

7.8. Co-localization between proteins : testing procedure and generative models

Participants: Frédéric Lavancier, Thierry Pécot, Charles Kervrann.

In the context of bioimaging, co-localization refers to the detection of emissions from two or more fluorescent molecules within the same pixel of the image. This approach enables to quantify the protein-protein interactions inside the cell, just at the resolution limit of the microscope. In statistics, this amounts to characterizing the joint spatial repartition and the spatial overlap between different fluorescent labels. An illustration of the co-localization of green (Langerin protein) and red (Rab11 GTPase protein) fluorescence is shown in Fig. 9 (the images were segmented by applying the ATLAS algorithm [12]). In our framework, the spatial repartition of proteins in the same cell is modeled by a union of random balls, possibly overlapping, and a Gibbs interaction is introduced to take into account the possible interaction between the two co-expressed proteins. A simulation algorithm is described and an inference procedure, based on the Takacs-Fiksel method, is proposed to estimate the interaction parameter. This estimation allows us to determine the presence of co-localization and to quantify the degree of interactions. On the other hand, this model can be used as a generator for synthesized images of co-localized proteins, in a view to assess testing procedures as the one explained below.

In an on-going project, we are developing a non-parametric testing procedure for co-localization. It is mainly based on the overlap area, corresponding to yellow spots as displayed in the right-hand side image of Fig. 9. Our first experiments on synthesized images showed that our procedure is more powerful than all existing methods to detect co-localization. Moreover this testing procedure turns out to be robust to different shapes and sizes of objects segmented by any competitive algorithm.

Reference: [36]



Figure 9. M10 cell showing Langerin proteins (left, in green) and Rab11 GTPase proteins (middle, in red). Right: superposition of the two previous images resulting in some possible yellow spots (co-expression of proteins within the same pixel).

7.9. Classification of diffusion dynamics from particle trajectories

Participants: Vincent Briane, Charles Kervrann.

In this study, we are currently interested in describing the dynamics of particles inside live cell. We assume that the motions of particles follow a certain class of random process: the diffusion processes. We have proposed a statistical method able to classify the motion of the observed trajectories into three groups: “confined”, “directed” and “free diffusion” (namely Brownian motion). This method is an alternative to the commonly used Mean Square Displacement (MSD) analysis. We assessed our procedure on both simulations and real cases; an example of confined diffusion is the Ornstein-Uhlenbeck process while an example of directed diffusion is the Brownian motion with constant drift. The method is currently applied to investigate membrane trafficking (Rab11/Langerin (see Fig. 10) and Rab11/TfR protein sequences) using the following procedure:

1. Tracking of particles with any competitive algorithm.
2. Statistical test /classification applied on tracks longer than ten time points.
3. Estimation of diffusion parameters (e.g. drift, diffusion, ...).

Each trajectory is labelled with the most likely process and the parameters of the underlying process are estimated. Future work will concern the detection of change of motion dynamic over time. Some results of our test on the Langerin protein sequence are shown in Fig. 10 .

Collaborator: Myriam Vimond (ENSAI Rennes).

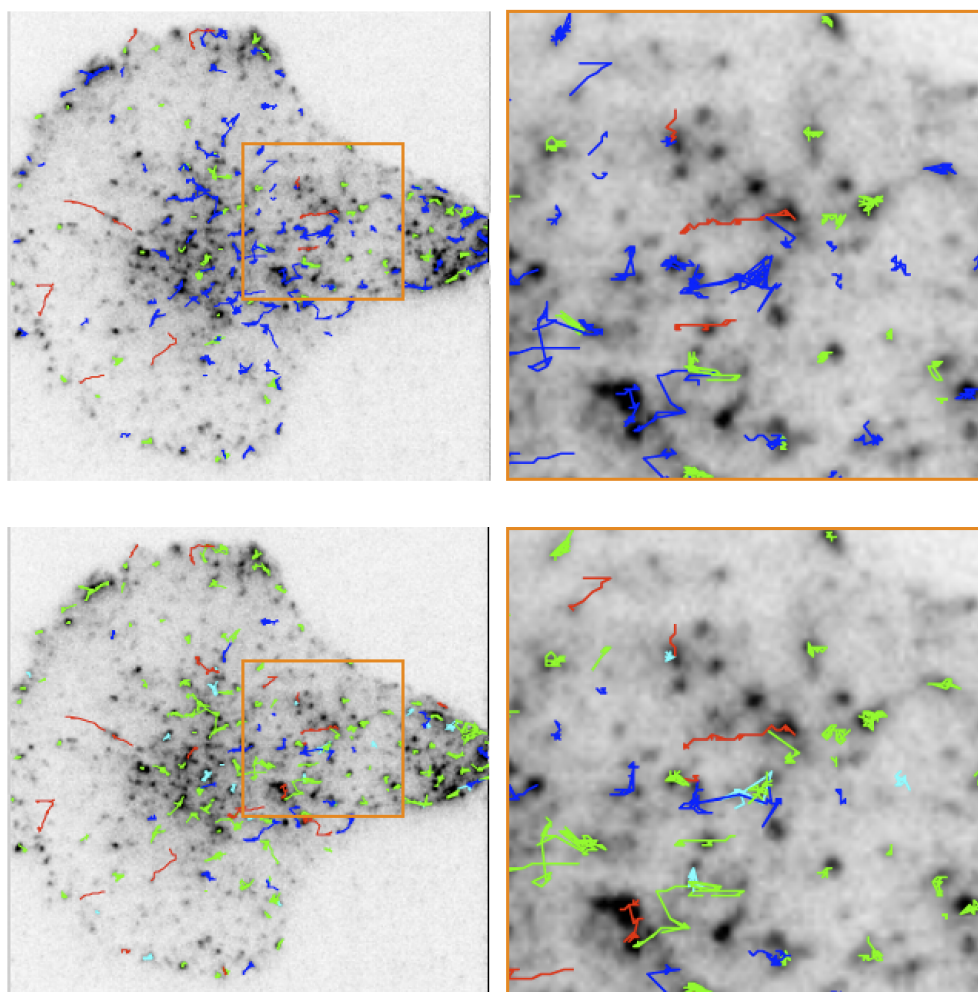


Figure 10. Labelling of the dynamics of trajectories on the Langerin protein sequence (Courtesy of UMR 144 CNRS-Institut Curie and PICT IBiSA). We display only the trajectories appearing on the first 100 frames. The color code is red for directed Brownian, green for Ornstein-Uhlenbeck, blue for Brownian, cyan for motionless. Top panel is labelled with our test, bottom panel with the MSD method.

7.10. Inference for spatial Gibbs point processes

Participant: Frédéric Lavancier.

Gibbs point processes are popular and widely used models in spatial statistics to describe the repartition of points or geometrical structures in space. They initially arose from statistical physics where they are models for interacting particles. They are now used in as different domains as astronomy, biology, computer science, ecology, forestry, image analysis and materials science. Assuming a parametric form of the Gibbs interaction, the natural method to estimate the parameters is likelihood inference. Since its first use in the 80's, this method is conjectured to be consistent and efficient. However the theoretical properties of maximum likelihood for Gibbs point processes remain largely unknown. In [39], we have partly solved this 30 years old conjecture by proving the consistency of the likelihood procedure for a large class of Gibbs models. As important examples, we deduced the consistency of the maximum likelihood estimator for all parameters of the Strauss model, the hardcore Strauss model, the Lennard-Jones model and the area-interaction model, which are commonly used models in practice.

A practical issue of likelihood estimation yet is that this method depends on an intractable normalizing constant that has to be approximated by simulation. To avoid this problem, other methods of estimation have been introduced, including pseudo-likelihood estimation. The theoretical properties of the pseudo-likelihood method are fairly well known in the case of finite-range Gibbs interactions. However, this setting rules out some major Gibbs models as the Lennard-Jones model. In [15], we have extended the pseudo-likelihood procedure to infinite range Gibbs interactions and proved its consistency and its asymptotic normality.

References: [15], [39]

Collaborators: David Dereudre (Laboratoire Paul Painlevé (UMR 8524), University of Lille 1),
Jean-François Coeurjolly (Laboratoire Jean Kutzmann, University of Grenoble).

7.11. Statistical aspects of Determinantal Point Processes

Participant: Frédéric Lavancier.

Determinantal point processes (DPPs) have been introduced in their general form by Macchi (1975) and have been extensively studied from a probabilistic point of view in the 2000's (one of the main reason being their central role in random matrix theory). In [23], we have demonstrated that DPPs provide useful models for the description of spatial point pattern datasets where nearby points repel each other. We have exploited the appealing probabilistic properties of DPPs to develop parametric models, where the likelihood and moment expressions can be easily evaluated and realizations can be quickly simulated. We have discussed how statistical inference is conducted using the likelihood or moment properties of DPP models, and we provided freely available software for simulation and statistical inference.

In [13], we have addressed the question of how repulsive a stationary DPP can be, in order to assess the range of practical situations this promising class of models may model. We determine the most repulsive DPP (in some sense) and we introduce new parametric families of stationary DPPs that can cover a large range of DPPs, from the stationary Poisson process (the case of no interaction) to the most repulsive DPP. Some theoretical aspects of inference for stationary DPPs are tackled in [37] and [38]. In the former study we have established the Brillinger mixing property of stationary DPPs, a first important step toward asymptotic inference. In the latter contribution, we have exploited this result to deduce the consistency and asymptotic properties of contrast estimators for stationary DPPs.

References: [23], [13], [37], [38]

Collaborators: Christophe Ange Napoléon Biscio (LMJL, University of Nantes),
Jesper Møller (Department of Mathematical Sciences, Aalborg University, Denmark),
Ege Rubak (Department of Mathematical Sciences, Aalborg University, Denmark).

7.12. Modelling aggregation and regularity in spatial point pattern datasets

Participant: Frédéric Lavancier.

In the spatial point process literature, analysis of spatial point pattern datasets are often classified into three main cases: i/ regularity (or inhibition or repulsiveness), modelled by Gibbs point processes, hard core processes like Matern hard core models, and determinantal point processes; ii/ complete spatial randomness, modelled by Poisson point processes; iii/ aggregation (or clustering), modelled by Poisson cluster processes and Cox processes. For applications the classification i/-iii/ can be too simplistic, and there is a lack of useful spatial point process models with, loosely speaking, aggregation on the large scale and regularity on the small scale. For instance, we may be interested in such a model for the repartition of the centres of vesicles in a cell, that exhibit some spatial clustering at large scales while having a minimal distance between them.

In [22], we have considered a dependent thinning of a regular point process with the aim of obtaining aggregation on the large scale and regularity on the small scale in the resulting target point process of retained points. Various parametric models for the underlying processes are suggested and the properties of the target point process are studied. Simulation and inference procedures have been discussed when a realization of the target point process is observed, depending on whether the thinned points are also observed or not. Some typical simulations of the target processes are shown in Fig. 11 .

Reference: [22]

Collaborator: Jesper Møller (Department of Mathematical Sciences, Aalborg University, Denmark).

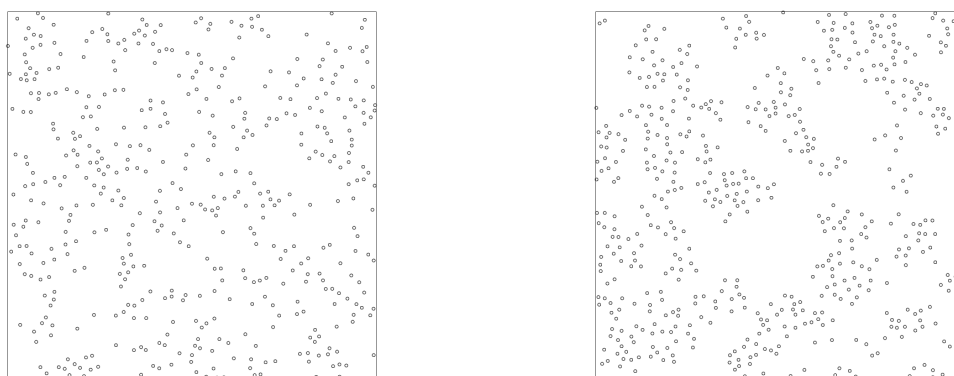


Figure 11. Examples of simulations with aggregation on the large scale and regularity on the small scale.

7.13. Retracing and registration for Correlative light-electron microscopy (CLEM)

Participants: Bertha Mayela Toledo Acosta, Patrick Boutheymy, Charles Kervrann.

Correlative light-electron microscopy (CLEM) enables to relate cell dynamics visualized in light microscopy (LM) with cell structure provided by electron microscopy (EM) for a better understanding of cell mechanisms. Registration of LM and EM modalities is then a timely, important but difficult open problem, which still requires some manual assistance. LM and EM images are indeed of very different size, spatial resolution, field of view, and appearance. We have investigated an original automated approach for the retracing-and-registration stage of the overall CLEM workflow (see Fig. 12). Pairing between the LM region of interest (ROI) and the corresponding EM patch relies on a common representation for both images, based on the LoG (Laplacian of Gaussian) transform with an adaptive associated scale (or blurring). We exploit histograms of the LoG values or histograms supplied by the LDP (Local Directional Pattern) texture descriptor, with associated histogram distances, to solve the EM patch search issue. The search step supplies a pre-registration, which is

refined by the estimation of an affine motion model to overlay the EM image onto the LM image around the ROI. Preliminary results on real CLEM images provided by UMR 144 CNRS-Institut Curie demonstrated the interest and efficiency of the proposed method.

Collaborators: Perrine Paul-Gilloteaux and Xavier Heiligenstein (UMR 144 CNRS-Institut Curie).

7.14. Denoising and compensation of the missing wedge in cryo electron tomography

Participants: Emmanuel Moebel, Charles Kervrann.

In this study, we have addressed two important issues in cryo electron tomography (CET) images: the low signal-to-noise ratio and the presence of a missing wedge (MW) of information in the spectral domain. Indeed, according to the Fourier slice theorem, limited angle tomography results into an incomplete sampling of the Fourier domain. Therefore, the Fourier domain is separated into two regions: the known spectrum (KS) and the unknown spectrum, the latter having the shape of a missing wedge (see Fig. 13). The proposed method tackles both issues jointly, by iteratively applying a denoising algorithm in order to fill up the MW, and proceeds as follows:

1. Excitation step: Add noise into the MW.
2. Denoising step: Apply a patch-based denoising algorithm.
3. Repeat steps 1 and 2, by keeping KS constant through the iterations.

The excitation step is used to randomly initialize the coefficients of the MW, whereas the denoising step acts as a spatial regularization. The employed denoising algorithm, which exploits the self-similarity of the image, filters out coefficient values which are dissimilar to KS, thereby keeping similar ones. By iterating these steps, we are able to diffuse the information contained in KS into the MW.

An application example on experimental data can be seen on Fig. 13, which shows the data in both spectral and spatial domain. The data contains a spherical gold particle, deformed by MW induced artifacts: elongation of the object, side- and ray-artifacts. From the residue image it can be seen that noise and MW artifacts have been reduced, while preserving the details of the image. Experiments are being performed to verify if particle detection and alignment are enhanced by using the method as a pre-processing step.

Collaborators: Damien Larivière (Fondation Fourmentin-Guilbert),
Julio Ortiz (Max-Planck Institute, Martinsried, Germany).

7.15. Algorithms for row registration to improve quality of Tissue MicroArray (TMA) images

Participants: Hoai Nam Nguyen, Charles Kervrann.

Row jittering is a common problem arising in medical imaging devices such as CT (Computer Tomography) and MRI (Magnetic Resonance Imaging) scanners due to errors of synchronization during image acquisition process. On scanners designed and developed by Innopsys, the problem becomes more challenging mainly because the pixel displacement is non constant along each row (Fig. 14) and possibly sub-pixel (i.e. non integer translation). To overcome this drawback, we first proposed a window-based algorithm to approximate the translation at each pixel by selecting the value that best minimizes a matching criteria over a finite set of possible sub-pixel translations. We obtained satisfying results with this method on real data with fast computation time (see Fig. 14). Furthermore, this matching criteria has been considered as a data fidelity term and was combined to a regularization term to promote a smooth solution and correct small artifacts which were not removed with the window-based method. To minimize the energy functional, we have adopted the quadratic relaxation technique and proximal method. This algorithm is slower and is initialized by the window-based algorithm to produce very encouraging results and elimination of all undesirable artifacts (see Fig. 14).

Collaborators: Vincent Paveau and Cyril Cauchois (Innopys).

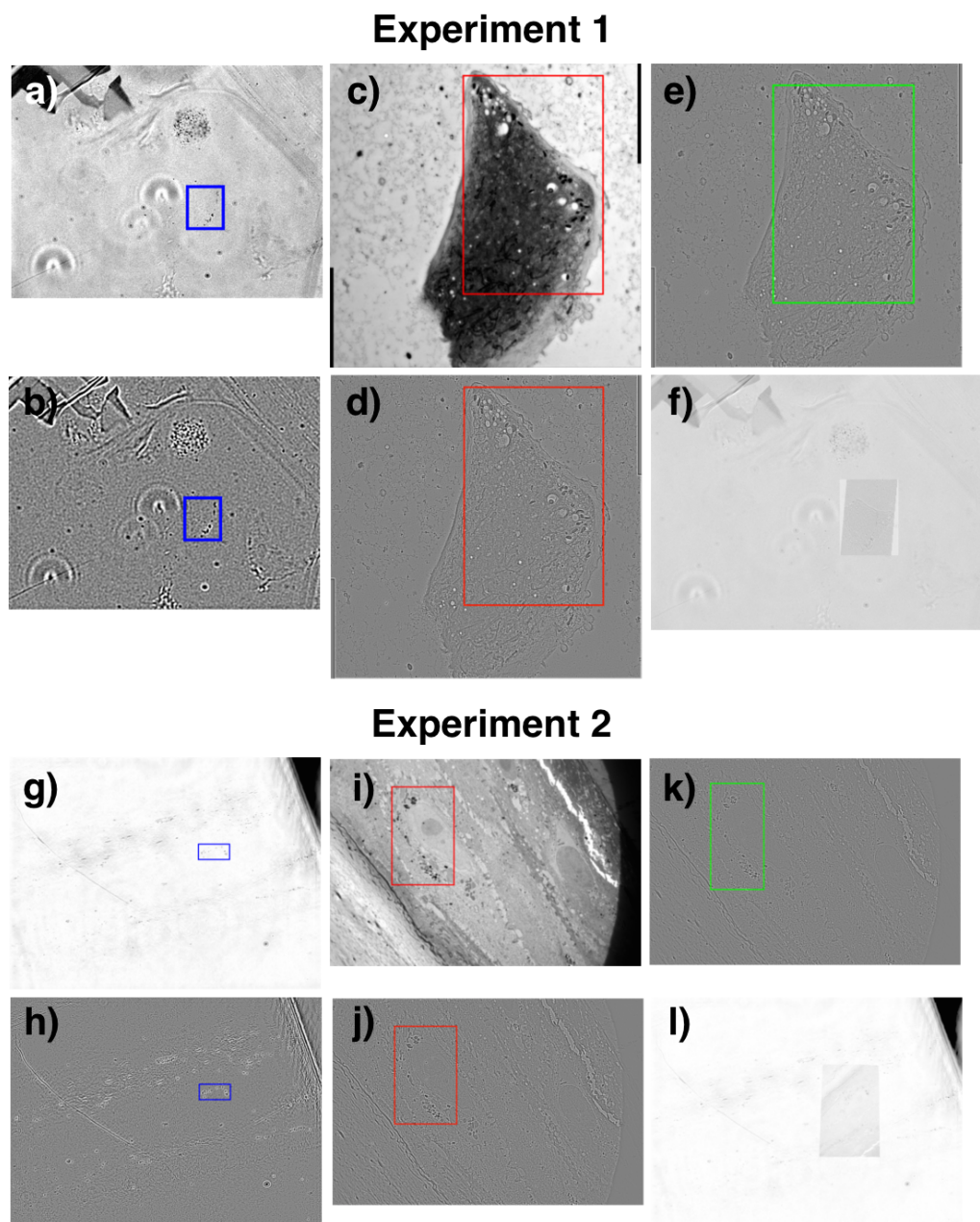


Figure 12. CLEM experiment #1: a) LM image with Region of Interest (ROI) framed in blue; b) same ROI delineated in the LoG-LM image; c) ground-truth location of the corresponding EM patch framed in red; d) the same but in the LoG-EM image; e) selected patch (SP) in the LoG-EM image in green; f) overlay (after registration) of the (decimated) EM image on the LM image around the ROI. CLEM experiment #2: g) LM-ROI in blue; h) LoG-LM-ROI; i) EM-GT in red; j) LoG-EM-GT; k) LoG-EM selected patch; l) Overlay of EM on LM around ROI.

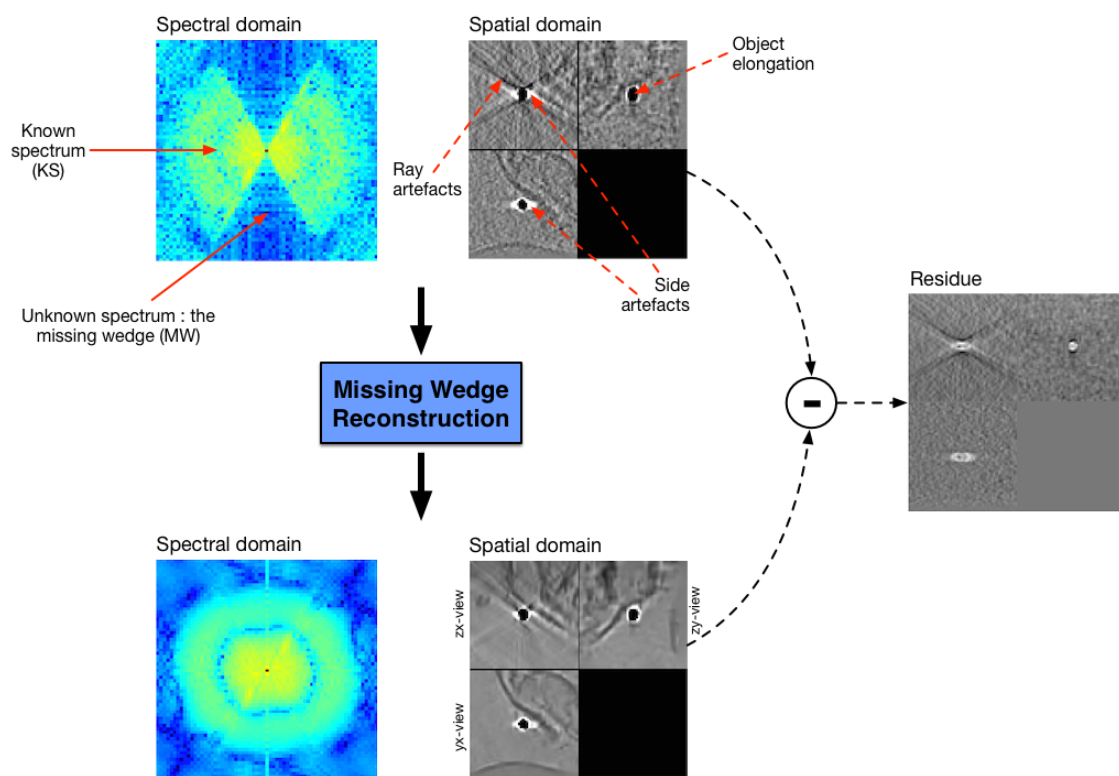


Figure 13. Experimental result of denoising and compensation of the missing wedge in cryo electron tomography.

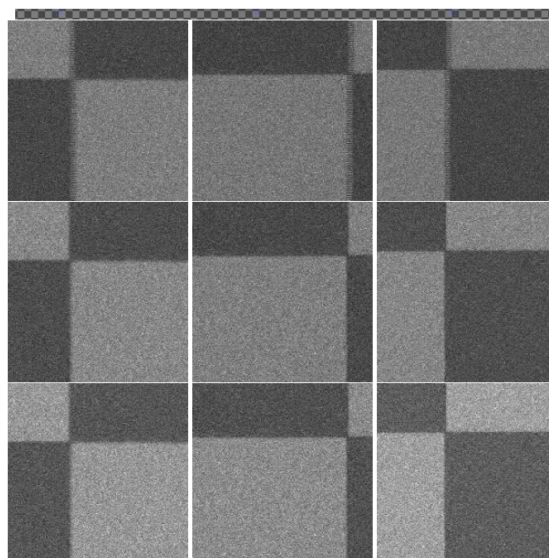


Figure 14. Illustration of the two-row registration algorithms. First row : full width input image (by courtesy of Innopsys). Second row : zooms on input image (blue boxes). Third row : corrected with window-based algorithm. Fourth row : corrected with variational method.

7.16. Robust motion model selection

Participants: Patrick Boutheymy, Bertha Mayela Toledo Acosta.

Parametric motion models are commonly used in image sequence analysis for different tasks. A robust estimation framework is usually required to reliably compute the motion model. However, choosing the most appropriate model in that estimation context is still an open issue. Indeed, penalizing the model complexity while maximizing the size of the inlier set may be contradictory. In this study, we proposed a robust motion model selection method which relies on the Fisher statistic. We also derived an interpretation of it as a robust C_P -Mallows criterion. The resulting criterion is straightforward to compute and explicitly involves the aforementioned trade-off between maximizing the size of the inlier set and minimizing the complexity (i.e., the number of parameters) of the selected motion model. We have conducted a comparative experimental evaluation on synthetic and real image sequences demonstrating that our criterion outperforms the RBIC criterion.

Collaborator: Bernard Delyon (IRMAR Rennes).

7.17. Anomaly detection in crowded scenes

Participants: Juan Manuel Perez Rua, Antoine Basset, Patrick Boutheymy.

We have defined an original motion-based method to detect and localize abnormal events in videos of crowded scenes. The algorithm relies on so-called labeled affine flows, involving both affine motion types and affine velocity vectors, and on view-based crowd motion classes. At every pixel the crowd motion class is inferred from the affine motion model selected among a set of candidate models estimated over a collection of windows. Then, the image is subdivided in blocks where local crowd motion class histograms weighted by the affine motion vector magnitudes are computed. They are block-wise compared to histograms of normal behaviors with a combined distance. More specifically, we introduce the so-called local outlier factor (LOF) to detect anomalous blocks. LOF is a local flexible measure of the relative density of data points in a feature space,

here the space of crowd motion class histograms. By thresholding the LOF value, we can detect an abnormal event in a given block at a given time. Comparative experiments on several real datasets demonstrated that our method is competitive with methods relying on far more elaborated models and exploiting both appearance and motion, while yielding superior performance over motion-based anomaly detection methods.

7.18. Occlusion detection in image sequences

Participants: Juan Manuel Perez Rua, Patrick Boutheymy.

The problem of localizing occlusions between consecutive frames of a video is important but rarely tackled on its own. In most works, it is tightly interleaved with the computation of accurate optical flows, which leads to a delicate chicken-and-egg problem. With this in mind, we proposed a novel approach to occlusion detection where visibility or not of a point in next frame is formulated in terms of visual reconstruction. The key issue is now to determine how well a pixel in the first image can be “reconstructed” from co-located colors in the next image. We first exploited this reasoning at the pixel level with a new detection criterion. Contrary to the ubiquitous displaced-frame-difference, the proposed alternative does not critically depend on a pre-computed, dense displacement field, while being shown to be more effective. We then leveraged this local modeling within an energy-minimization framework that delivers occlusion maps. An easy-to-obtain collection of parametric motion models is exploited within the energy to provide the required level of motion information. Our approach outperforms state-of-the-art detection methods on the challenging MPI Sintel dataset.

Collaborators: Tomas Crivelli and Patrick Pérez (Technicolor).

SISTM Project-Team

7. New Results

7.1. Time-Course Gene Set Analysis for Longitudinal Gene Expression Data

The application of TcGSA methodology has revealed the commitment of inflammatory pathways and T-cell pathway in response of DC-based vaccine.

STEPP Project-Team

6. New Results

6.1. Methods for the calibration of LUTI models

The setting up of a LUTI model requires, like most numerical models, at least one phase of parameter estimation. This is concisely referred to here as calibration, although the calibration of a LUTI model also entails other aspects such as the definition of spatial zones, of economic sectors, etc. The TRANUS LUTI model plus software, like many other existing models, come along with a relatively simple calibration methodology. Most LUTI models indeed perform parameter estimation in a piecewise fashion, by sequentially estimating subsets of parameters. While this reduces the mathematical and computational complexity of calibration, neglecting the interactions across different modules and their parameters, may result in a significant loss of a model's quality. A second issue is that TRANUS, like several other LUTI softwares, employs rudimentary numerical routines for parameter estimation. We aim at reducing these weaknesses.

In 2014, we had obtained first results along these lines: parameter estimation of the so-called shadow prices (specific parameters of the TRANUS model) was posed as optimization problem and several solution procedures were developed which were based on “unwinding” the dynamics of the model, making the problem amenable to standard numerical optimisation techniques.

The work continued throughout 2015, along different directions. First, the calibration was extended to handle several different parameter types simultaneously (shadow prices as well as the so-called substitution parameters, which are notoriously difficult to estimate) [7]. Such a simultaneous estimation of different parameter sets seems to be rare in LUTI practice.

Second, we proposed a methodology for assessing properties (convergence, accuracy) of our (and other) LUTI calibration methods [6]. This consists in generating synthetic data, starting from a model calibrated on observed data, such that the synthetic data are completely consistent, i.e. there are a set of model parameters that exactly reproduce these data (which is not the case with the observed data). The ground truth model parameters are then easily used to assess calibration parameters. Such a methodology, akin to twin experiments in data assimilation, seems to be novel for LUTI research.

Third, LUTI models are usually calibrated on a base year or period, and used in a prospective manner (via simulated “predictions” for future periods). As with any numerical model, it is wise to make sure that a calibrated LUTI model does not overfit the observations used for calibration; otherwise, its “predictions” may be grossly erroneous. Potential overfitting does not seem to have been deeply studied in the LUTI literature. We have made an initial investigation by calibrating different versions of a TRANUS model, varying the number of shadow prices used as parameters in the model (there is, by default, one shadow price per combination of geographical zone of the study area and economic sector) [6]. For instance, after an initial calibration using all shadow prices, we then dropped the two third smallest of them and re-calibrated the model using the remaining third. The goodness-of-fit to observations was worse by only 3%. In line with well-known principles of model selection (Occam's razor), this may suggest that it is preferable to use the model with fewer parameters when doing predictions. This is still work in progress; showing its relevance is planned to be studied by a similar methodology as above, using simulated twin experiments.

This work is done in collaboration with Arthur Vidard from the AIRSEA Inria project-team and Brian Morton from the University of North Carolina at Chapel Hill.

6.2. Estimation of Sobol' indices combining nested designs and replication method

Sensitivity analysis studies how the uncertainty on an output of a mathematical model can be attributed to sources of uncertainty among the inputs. Global sensitivity analysis of complex and expensive mathematical

models is a common practice to identify influent inputs and detect the potential interactions between them. Among the large number of available approaches, the variance-based method introduced by Sobol' allows to calculate sensitivity indices called Sobol' indices. Each index gives an estimation of the influence of an individual input or a group of inputs. These indices give an estimation of how the output uncertainty can be apportioned to the uncertainty in the inputs. One can distinguish first-order indices that estimate the main effect from each input or group of inputs from higher-order indices that estimate the corresponding order of interactions between inputs. This estimation procedure requires a significant number of model runs, number that has a polynomial growth rate with respect to the input space dimension. This cost can be prohibitive for time consuming models and only a few number of runs is not enough to retrieve accurate informations about the model inputs.

The use of replicated designs to estimate first-order Sobol' indices has the major advantage of reducing drastically the estimation cost as the number of runs becomes independent of the input space dimension. The generalization to closed second-order Sobol' indices relies on the replication of randomized orthogonal arrays. However the replication method still requires a large number of model evaluations. By rendering this method iterative, the required number of evaluations can be controlled. The estimation procedure is therefore stopped when the convergence of estimates is considered reached. The key feature of this approach is the construction of nested designs. For the estimation of first-order indices, we exploit a nested Latin Hypercube already introduced in the literature. For the estimation of closed second-order indices, two methods are proposed to construct a nested orthogonal array. One of the two leads to a partition of the coordinate space over a Galois field.

This work has been done in collaboration with Laurent Gilquin and Clementine Prieur (members of Moise Team), and belongs to the work program of CiTIES project. The proposed procedure will be soon applied to study the sensitivity of TRANUS model.

6.3. Environmental pressures associated with material flows

This work is the follow-up of a previous study dedicated to material flow analysis of the French cereal supply chain at various spatial levels [12]. The goal was twofold:

- trace the flows to their initial geographic origin or final destination,
- couple material flows with a series of environmental pressures associated to them.

For the first goal, we used an Absorbing Markov Chains model where transient states represent raw or semi-products and absorbing states correspond to final consumption products. For the second goal, we used pressure ratios for environmental pressures most relevant to cereals, namely energy use, GHG emissions, land use, use of pesticides and blue water footprint. The model is based on physical supply and use tables and distinguishes between 21 industries, 22 products, 38 regions of various spatial resolution (22 French regions, 10 countries, 6 continents) and 4 modes of transport. Illustrative examples were taken in order to demonstrate the versatility of the results produced, for instance: What is the fate/supply area of a region's production/consumption? What are the production and consumption footprint of a region? These results are designed to be a first step towards scenario analysis for decision-aiding that would also include socioeconomic indicators [13]

6.4. Material flows of the French forest-wood supply chain

The methodology developed in Courtonne et al. [12] on the case of the cereal supply chain was adapted to the French forest-wood supply chain in collaboration with the Laboratoire d'Economie Forestière. Supply chain flows were estimated both at the national and regional scale for wood harvest, addition to stock, production, imports and exports of construction wood, industrial wood and energy wood. These results can be a basis to analyze potential value losses throughout the supply chain, for instance exports of raw materials instead of local transformation. They can also be use to study the competitive use of wood for energy, industry and construction/furnitures, which is a question of growing importance in the context of energy transition.

6.5. Land Use/Land Cover Change (LUCC) Modelling and Ecosystem Services

The ESNET project (EcoSystem services NETworks) is a collaboration lead by LECA (Laboratoire d'ECologie Alpine, UJF) that aims at characterizing the ecosystem services of the Grenoble urban region (about 2/3 of the Isere département) at the 2040 horizon under various constraints of urban policy planning, changes in agricultural and forest management, and climate change impact on ecosystems.

The cartographic effort of the project has been hosted at Inria, and has produced in 2014 three very detailed maps of land use and land cover at the 15m resolution over the whole study area, in 1998, 2003 and 2009, respectively. An extensive analysis of the patterns of landscape change has been performed from these data, with special emphasis on urban sprawl and the associated loss of arable land. This work has been submitted for publication very recently.

A second related piece of work has been produced, both from this cartographic source and more specific remote sensing data. The objective was to characterize in detail the cultural successions and patterns of the study area, in order to produce fine scale maps of associated ecosystem services. This work has just been submitted for publication at the time of writing.

Finally, the scenarios of future land use and land cover that have been elaborated for this project have all been projected at the 2040 horizon at the 15m scale with a well-known LUCC modelling environment (Dinamica) for urban changes, and from in-project models for the other types of land use and cover. A third article bearing on these scenarios and their LUCC modelling is in preparation.

As an aside of this land use/cover modelling effort, the STEEP team has been involved in two of the most detailed ecosystem service models developed for the project: one for the analysis of crop production and associated nitrogen cycle assessment — with the final aim to constrain both the production services and water quality issue related to nitrogen loading — and one on “recreational services”. Our involvement in these models was directly related to the acquired expertise in land use modelling.

In the process of this modelling exercise, the STEEP team has acquired an in-depth knowledge and expertise of LUCC models. As a consequence, various theoretical flaws have been identified in the theoretical foundations of such models. An important by-product of the ESNET project is therefore a series of articles in preparation in the team, whose aim is to address and correct these flaws in a very general way; it is hoped that LUCC theory will be put on a more serious theoretical footing as a result of this methodological work, which should be submitted for publication in 2016 for the most part. Another but more limited methodological contribution bears on the development of error models for landscape metrics, another important methodological blind-spot in the specialized literature.

6.6. A benchmarking tool to assess the compatibility of the INDCs with the 2°C long-term target

Climate negotiations related to global warming are another important issue of sustainable development. In this framework that is place at international scale we have developed a benchmarking tool which allows to assess the compatibility of the Intended Nationally-Determined Contributions (INDCs) given by all states for the Conference COP21, with the 2°C long-term target. This benchmarking tool has been designed via an adaptation of REDEM model and algorithm we developed in 2014 with EDDEN laboratory. This tool has been used by the “*Groupe Interdisciplinaire sur les Contributions Nationales*” (GICN) which has been mandated by french ministry of Sustainable Development to prepare the climate change conference COP21 at Paris.

TONUS Team

7. New Results

7.1. Particle-in-cell simulations for highly oscillatory Vlasov-Poisson systems

Participants: Edwin Chacon Golcher, Sever Adrian Hirstoaga [correspondent], Mathieu Lutz.

The aim of the following works is to study the dynamics of charged particles under the influence of a strong magnetic field by numerically solving in an efficient way the Vlasov-Poisson and guiding center models.

First, we work on the development of the time-stepping method introduced in [7], [8] in two directions: improve the accuracy of the algorithm and adapt the algorithm for general configuration of magnetic field.

Second, by using appropriate data structures, we implement an efficient (from the memory access point of view) Particle-In-Cell method which enables simulations with a large number of particles. Thus, we present in [13] numerical results for classical one-dimensional Landau damping and two-dimensional Kelvin-Helmholtz test cases. The implementation also relies on a standard hybrid MPI/OpenMP parallelization. Code performance is assessed by the observed speedup and attained memory bandwidth. A convergence result is also illustrated by comparing the numerical solution of a four-dimensional Vlasov-Poisson system against the one for the guiding center model.

7.2. Eulerian simulations of parallel transport in the SOL

Participants: David Coulette, Sever Adrian Hirstoaga [correspondent], Giovanni Manfredi.

We continue to investigate kinetic models for simulating the heat load on the divertor plates during transient events as edge-localised modes (ELMs). Our previous work [36] deals with Vlasov-Poisson equations for two particle species for the dynamics of their transport parallel to the magnetic field. We started to improve this model by adding an equation for the evolution in time of the perpendicular temperatures. These equations take also into account the collisions between species which may play a role over long times. The first numerical results are encouraging, showing different features with respect to the older (simpler) model when computing total particles and energy fluxes on the divertor plates.

7.3. Quasi-neutrality equation in a polar mesh

Participants: Christophe Steiner [correspondent], Michel Mehrenberger, Nicolas Crouseilles, Philippe Helluy.

In this work [21], we are concerned with the numerical resolution of the quasi-neutrality equation arising in plasma physics. A classic method is based on a Padé approximation. Two other methods are proposed in this paper: a high order Padé approximation and a direct method in the space configuration which consists in integrating on the gyrocircles using an interpolation operator. Numerical comparisons are performed with analytical solutions and considering the 4D drift-kinetic model with one Larmor radius. This is a preliminary study; further study in GYSELA is envisioned.

7.4. The Semi-Lagrangian method on curvilinear grids

Participants: Aurore Back, Adnane Hamiaz, Michel Mehrenberger [correspondent], Pierre Navaro, Hocine Sellama, Eric Sonnendrücker.

We study the semi-Lagrangian method on curvilinear grids [18], [9]. The classical backward semi-Lagrangian method preserves constant states but is not mass conservative. Natural reconstruction of the field permits nevertheless to have at least first order in time conservation of mass, even if the spatial error is large. Interpolation is performed with classical cubic splines and also cubic Hermite interpolation with arbitrary reconstruction order of the derivatives. High odd order reconstruction of the derivatives is shown to be a good ersatz of cubic splines which do not behave very well as time step tends to zero. A conservative semi-Lagrangian scheme is then described; here conservation of mass is automatically satisfied and constant states are shown to be preserved up to first order in time.

Semi-Lagrangian guiding center simulations are performed on sinusoidal perturbations of cartesian grids, and on deformed polar grids with different boundary conditions. Key ingredients are: the use of a B-spline finite element solver for the Poisson equation and the classical backward semi-Lagrangian method (BSL) for the advection. We are able to reproduce standard Kelvin-Helmholtz and diocotron instability tests on such grids. When the perturbation leads to a strong distorted mesh, we observe that the solution differs if one takes standard numerical parameters that are used in the cartesian reference case. We can recover good results together with correct mass conservation, by diminishing the time step.

7.5. Solving the Guiding-Center model on a regular hexagonal mesh

Participants: Michel Mehrenberger [correspondent], Laura Mendoza, Charles Prouveur, Eric Sonnendrücker.

This work [11] introduces a Semi-Lagrangian solver for the Vlasov-Poisson equations on a uniform hexagonal mesh. The latter is composed of equilateral triangles, thus it doesn't contain any singularities, unlike polar meshes. We focus on the guiding-center model, for which we need to develop a Poisson solver for the hexagonal mesh in addition to the Vlasov solver. For the interpolation step of the Semi-Lagrangian scheme, a comparison is made between the use of box-splines and of Hermite finite elements. The code will be adapted to more complex models and geometries in the future.

7.6. High-order Hamiltonian splitting for Vlasov-Poisson equations

Participants: Fernando Casas, Nicolas Crouseilles, Erwan Faou, Michel Mehrenberger [correspondent].

In this work [12], we consider the Vlasov-Poisson equation in a Hamiltonian framework and derive new time splitting methods based on the decomposition of the Hamiltonian functional between the kinetic and electric energy. Assuming smoothness of the solutions, we study the order conditions of such methods. It appears that these conditions are of Runge-Kutta-Nyström type. In the one dimensional case, the order conditions can be further simplified, and efficient methods of order 6 with a reduced number of stages can be constructed. In the general case, high-order methods can also be constructed using explicit computations of commutators. Numerical results are performed and show the benefit of using high-order splitting schemes in that context. Complete and self-contained proofs of convergence results and rigorous error estimates are also given.

7.7. Velocity space transformations: collisional case

Participants: Emmanuel Franck, Philippe Helluy [correspondent], Laurent Navoret.

The method of "velocity space transformations" allows to obtain an interesting discretization of the Kinetic equations like Vlasov-Poisson or Vlasov Maxwell equations as has been proved in the works of P. Helluy, L. Navoret and N. Pham. During this year, we have begun to extend this method to the collisional case using the entropy variable to write a general collisional operator. To treat all the regimes (small or large collisional regime), asymptotic preserving schemes (stability and convergence independent of the collisional frequency) have been designed. However, this method admits some numerical difficulties if we use the physical entropy to construct the collisional operator. Now we propose to use modified entropy, which has good numerical properties and gives limit regime close to the real one in the low Mach context. If this new approach gives interesting results, we will study the adaptivity of the velocity discrete basis which would allow to treat the collisional and non-collisional regimes with the same method.

7.8. Preconditioning and implicit solvers

Participants: Emmanuel Franck [correspondent], Philippe Helluy, Matthias Hoelzl, Ahmed Ratnani, Malcolm Roberts, Eric Sonnendrücker, Stefano Serra-Capizzano.

The Viscous-resistive MHD model used to simulate the instabilities is a multi-scale models with fast waves. In this context, it is not possible to use full explicit time schemes. However the classical implicit schemes are not usable directly since the matrices are ill-conditioned. For this reason it is necessary to use a preconditioning method. During this year we have studied a method called "physic based preconditioning" for the wave equations which consists to approximate the solution by suitable smaller and simpler systems. The results are very good. After this, we have extended this method to the Linearized Euler equation. During this new study, we have found additional difficulties which appear in some regimes. Two methods to treat this problem will be tested in 2016. We have also implemented a version of this preconditioning for the reduced MHD models of JOREK. The first results are positive. To finish, we have begun a collaboration with S. Serra-Capizzano to study at the theoretical level the physic based preconditioning and propose new preconditioning for each sub-systems of the Physic-Based PC efficient in all the physics regimes and for an arbitrary order.

We have also developed an implicit solver for the transport equation based on the upwind nature of the DG numerical flux. This solver will be used for solving Vlasov models or fluid models thanks to the Lattice-Boltzmann methodology. We have obtained recently a SPPEXA support (<http://www.sppexa.de>) in a joint french-german-japanese project.

7.9. Finite element for full-MHD problems

Participants: Emmanuel Franck [correspondent], Eric Sonnendrücker.

This work have begun at the end of 2015. It is organized around a PhD: Mustafa Gaja supervised by E. Sonnendrücker, A. Ratnani and E. Franck at the Max-Planck Institute of Plasma Physic. The aim of this work is to design and study compatible finite element method (finite element method which preserve the DeRham sequence and the inclusion between the functional space) for B-Splines. This method will allow to discretize efficiently the Maxwell equations, the MHD model and some operators as curl-curl or grad-div vectorial operators which appear in the physic-based PC. For now, we have begun to study the finite element discretization of vectorial operators which appears in the linearized Euler equations and in the physic-based PC associated.

7.10. Lagrangian averaged gyrokinetic-waterbag continuum

Participant: Nicolas Besse [correspondent].

In this paper [26], we first present the derivation of the anisotropic Lagrangian averaged gyrowaterbag continuum (LAGWBC- α) equations. The gyrowaterbag (nickname for gyrokinetic-waterbag) continuum can be viewed as a special class of exact weak solution of the gyrokinetic-Vlasov equation, allowing to reduce this latter into an infinite dimensional set of hydrodynamic equations while keeping its kinetic features such as Landau damping. In order to obtain the LAGWBC- α equations from the gyrowaterbag continuum we use an Eulerian variational principle and Lagrangian averaging techniques introduced by Holm, Marsden, Ratiu and Shkoller for the mean motion of ideal incompressible flows, extended to barotropic compressible flows by Bhat and co-workers and some supplementary approximations for the electrical potential fluctuations. Regarding to the original gyrowaterbag continuum, the LAGWBC- α equations show some additional properties and several advantages from the mathematical and physical viewpoints, which make this model a good candidate for describing accurately gyrokinetic turbulence in magnetically confined plasma. In the second part of this paper we prove local-in-time well-posedness of an approximate version of the anisotropic LAGWBC- α equations, that we call the "isotropic" LAGWBC- α equations, by using quasilinear PDE type methods and elliptic regularity estimates for several operators.

7.11. Hamiltonian structure, fluid representation, stability for the Vlasov-Dirac-Benney equation

Participants: Claude Bardos, Nicolas Besse [correspondent].

This contribution [23] is an element of a research program devoted to the analysis of a variant of the Vlasov–Poisson equation that we dubbed the Vlasov–Dirac–Benney equation or in short V–D–B equation. As such it contains both new results and efforts to synthesize previous observations. One of main links between the different issues is the use of the energy of the system. In some cases, such energy becomes a convex functional and allows to extend to the present problem the methods used in the study of conservation laws. Such use of the energy is closely related to the Hamiltonian structure of the problem.

7.12. Semi-classical limit of an infinite dimensional system of nonlinear Schrödinger equations

Participants: Claude Bardos, Nicolas Besse [correspondent].

In this paper [24], we study the semi-classical limit of an infinite dimensional system of coupled nonlinear Schrödinger equations towards exact weak solutions of the Vlasov-Dirac-Benney equation, for initial data with analytical regularity in space. After specifying the right analytic extension of the problem and solutions, the proof relies on a suitable version of the Cauchy-Kowalewski Theorem and energy estimates in Hardy type spaces with convenient analytic norms.

7.13. Aligned interpolation for gyrokinetic Tokamak simulations

Participants: Guillaume Latu, Michel Mehrenberger [correspondent], Maurizio Ottaviani, Eric Sonnendrücker.

This work is devoted to study the aligned interpolation method in semi-Lagrangian codes. The scheme is presented and algorithms used implementing the scheme are given. A theoretical justification of the method is given with convergence estimates in the simplified context of 2D constant advection, assuming stability of the scheme. The stability is here studied numerically, letting the formal proof as an open problem. The solution is successfully applied in the gyrokinetic context: first in a simplified case in cylindrical geometry and then in toroidal geometry. In the first case, the solutions provided by simulations based on the scheme are in accordance with linear dispersion analysis; in the second case, numerical simulations produced by the Gysela code are presented, simulation based on the standard scheme are compared to those based on the new aligned scheme. This work will lead to a project of paper, which will be submitted in 2016.

VIRTUAL PLANTS Project-Team

6. New Results

6.1. Analysis of structures resulting from meristem activity

6.1.1. Acquisition and design of plant geometry

Participants: Frédéric Boudon, Christophe Pradal, Christophe Godin, Christian Fournier, Ibrahim Chedaddi, Mathilde Balduzzi, Julien Diener.

Virtual 3D model of plants are required in many areas of plant modeling. They can be used for instance to simulate physical interaction of real plant structures with their environment (light, rain, wind, pests, ...), to set up initial conditions of growth models or to assess their output against real data. In the past decade, methods have been developed to digitize plant architectures in 3D [76], [63]. These methods are based on direct measurements of position and shape of every plant organ in space. Although they provide accurate results, they are particularly time consuming. More rapid and automated methods are now required in order to collect plant architecture data of various types and sizes in a systematic way. In this aim, we explore the use of laser scanner and pictures.

- *Reconstruction of tree structures from 3D laser scanner data.* (Chakkrit Preuksakarn, Mathilde Balduzzi, Frédéric Boudon, Christophe Godin, Pascal Ferraro [Labri, Bordeaux], Yassin Refahi)

We investigate the possibility to use 3D laser scanners to automate plant digitizing. We are developing algorithms to reconstruct branching systems without leaves or foliage from scanner data or from scan simulated on plant mock-up obtained using different digitizing method.

For the branching systems, we previously proposed a reconstruction method to reconstruct plausible branching structures from laser scanner data based on the concept of space colonization [73]. Additionally, a number of automatic methods were proposed in the literature. The question of their comparison and relative accuracy is however critical for further exploitation in biological applications. To address such problem, we developed an evaluation pipeline that takes two plant structures as input and compares their organization using two indices of geometrical and structural similarities [55]. A first comparative evaluation of the different methods of the literature has been designed and conducted. A graphical editor has been developed and makes it possible to test the different methods and correct manually the reconstruction. A procedure to automatically determine phyllotactic angles from scans of small plants has been added to the reconstruction pipeline and has been tested on database of 150 scans of *Arabidopsis thaliana* with different genotypes. The editor has also been tested on apple trees and large African trees.

In the context of the PhD of M. Balduzzi, we also investigated the reconstruction of tree foliage from 3D scans. Such elements are crucial to study the interaction of the plant with its environment. However, laser scans contain outliers in the silhouette of the scans that make the meshing of the pointset extremely difficult. New generation of laser scanners provide intensity of the laser reflected on the surface of scanned objects. This intensity depends on the distance to the object, its optical property and the incidence angle. A first work on this topic showed that after correcting the distance effect, the incidence angle can be deduced from the intensity. From this result, we developed a reconstruction pipeline using the scan intensities and based on Shape-From-Shading. Outliers being along the edge of the surface point cloud, we chose to develop a propagation SFS method initialized with points of the scans with high quality. We proved that surface with constant intensity are necessarily surfaces of constant slope or sand-pile surfaces. Using this result, a propagation method along iso-intensity regions was developed. These surfaces can then be sampled to provide a smooth point set without outliers.

- *Reconstruction of annual plants from multi-view images.* (Simon Artzet, Jerome Chopard, Christian Fournier, Christophe Pradal, Christophe Godin, Xavier Sirault [CSIRO-HRPPC, Canberra])

Image-based phenotyping platforms in semi-controlled conditions offer large possibilities to perform genetic analyses of plant growth, architecture, light interception, and biomass accumulation over large time series for thousands of plants. However, methods for image analyses currently available are still very crude and need improvement and robustness to process huge amount of data. We are developing an integrated pipeline allowing assessment of growths of individual organs, of plant geometry, and of derived variables such as light interception. The pipeline currently consists of 2D image analysis workflows built with standard image libraries (OpenCV, Scikit.Image), algorithms for 3D reconstruction, segmentation and tracking of plant organs for maize (under development), and workflows for estimation of light interception by plants during their growth. A 3D FSPM model for maize architectural development, is used to help segmenting plant images and to automate the mapping between segmented 3D objects and plant organs defined in the model.

- *Reconstruction of root structures.* (Julien Diener, Frédéric Boudon, Christophe Pradal, Christophe Godin, Philippe Nacry [BPMP, INRA], Christophe Périn [AGAP, CIRAD], Anne Dievart [AGAP, CIRAD], Xavier Draye [UCL, Belgium])

This research theme is supported by the Agropolis through the Rhizopolis project and by NUMEV.

Similarly to aerial part of plants, new needs for automatic digitizing of root systems emerge. Most existing methods focus only on semi-automatic approaches. This does not support the high-throughput capabilities of acquisition systems. In the context of the RhizoScan project, we previously designed a prototype of an automatic image analysis pipeline to extract root system architecture of branching systems grown in Petri boxes. This pipeline provides i) a set of model based image segmentation method, ii) the extraction of a graph representation of the root system, and iii) a method to identify the root axes organization. This year, we improved and extended the pipeline in the following way:

1. We integrated a validation step in the workflow based on the comparison method presented in [55].
 2. We developed a standard file format for root architecture (RSML) described in [19] during an international collaboration with the université Catholique de Louvain (Belgium), the CPIB of the University of Nottingham (UK), the University of Vienna (Austria), the Jülich research center (Germany) and INRA.
- *Reconstruction of virtual fruits from pictures.* (Ibrahim Chedaddi, Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2])

This research theme is supported by the Agropolis project MecaFruit3D.

The aim of this work is to provide methods for generating fruit structure that can be integrated with models of fruit function. To this end, a modeling pipeline has been developed in the OpenAlea platform. It involves two steps: (1) generating a 3D volumetric mesh representation of the entire fruit, and (2) generating a complex vascular network that is embedded within this mesh using the concept of space colonization [75]. Previous studies demonstrated the possibility to create species-specific models of fruit structure with relatively low effort [57]. We focus now on validating the vascular networks by comparing them to experimental data from the literature. This work has been presented at the ISHS symposium in Montpellier [38]

Using these fruit virtual structures, a mechanical model of fruit growth is also developed (see section 6.3.2) taking into account the distribution of water fluxes in the fruit.

6.1.2. Modeling the plant ontogenic programme

Participants: Christophe Godin, Yann Guédon, Jean-Baptiste Durand, Pierre Fernique, Marc Labadie, Christophe Pradal, Jean Peyhardi.

This research theme is supported by two PhD programmes.

The remarkable organization of plants at macroscopic scales may be used to infer particular aspects of meristem functioning. The fact that plants are made up of the repetition of many similar components at different scales, and the presence of morphological gradients, e.g. [52], [65], [66], [62], provides macroscopic evidence for the existence of regularities and identities in processes that drive meristem activity at microscopic scales. Different concepts have been proposed to explain these specific organizations such as "morphogenetic programme" [71], "age state" [61] or "physiological age" [54]. All these concepts state that meristem fate changes according to position within the plant structure and during its development. Even though these changes in meristem fate are specific to each species and lead to the differentiation of axes, general rules can be highlighted [61], [54]. Here we develop computational methods to decipher these rules.

- *Relating branching structure to the shoot properties* (Jean Peyhardi, Yann Guédon, Evelyne Coste [AGAP, AFEF team], Catherine Trottier [I3M], Yves Caraglio [AMAP], Pierre-Eric Lauri [AGAP, AFEF team])

Shoot branching structures often take the form of a succession of homogeneous branching zones and have been analyzed using segmentation models such as hidden semi-Markov chains. Axillary meristem fates are influenced by local properties of the parent shoot such as for instance its growth rate or local curvature. The objective of this work, which was part of the PhD subject of Jean Peyhardi, is to develop statistical models that generalize hidden semi-Markov chains with the capability to incorporate explanatory variables that vary along the parent shoot (e.g. leaf growth rate, leaf surface, internode length, local curvature of the parent shoot). More precisely, the simple multinomial distributions that represent the axillary productions observed in the different branching zones are replaced by multinomial generalized linear models (GLMs). Since the two classical categories of multinomial GLMs that correspond either to nominal or ordinal categorical response variables were not appropriate, we chose to develop a new family of multinomial GLMs called partitioned conditional GLMs [72] that enable to tackle hierarchically-structured categorical response variables. Typically, we need to distinguish different timing of branching events (e.g. immediate shoot, one-year-delayed shoot and latent bud), different categories of offspring shoots (e.g. among one-year-delayed shoots, vegetative short shoot, vegetative long shoot and flowering shoot) and to specialize the explanatory variables for certain categories of offspring shoots (e.g. the growth of the parent shoot influence the immediate offspring shoots but not the one-year-delayed offspring shoots). The resulting integrative models are called semi-Markov switching partitioned conditional GLMs and have been applied to apple and pear tree branching structures.

- *Genetic determinisms of the alternation of flowering in apple tree progenies*. (Jean-Baptiste Durand, Alix Allard [AGAP, AFEF team], Jean Peyhardi, Baptiste Guitton [AGAP, AFEF team], Yan Holtz [AGAP, AFEF team] Catherine Trottier, Evelyne Costes [AGAP, AFEF team], Yann Guédon)

A first study was published to characterize genetic determinisms of the alternation of flowering in apple tree progenies [58]. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot, which corresponds to portions of stem that were grown during the same year). Two replications of each genotype were available.

Indices were proposed for early detection of alternation during the juvenile phase. They were based on a trend model and a quantification of the deviation amplitudes and dependency, with respect to the trend. This allowed early quantification of alternation from the yearly numbers of inflorescences at tree scale. Some quantitative trait loci (QTL) were found in relation with this indices.

For better interpretation of the relationships of alternation at both scales, new models and indices were developed for sequences of flowering events at axis scale. New data sets were collected in other F1 progenies. Ancestral relationships between parents of different progenies were taken into account to enhance the power of QTL detection, and other QTL were found using these new indices.

- *Identifying and characterizing patterns in tree-structured data* (Pierre Fernique, Jean-Baptiste Durand, Yann Guédon).

In the context of Pierre Fernique's PhD (Montpellier 2 University and CIRAD), two complementary approaches were developed for analyzing patterns in tree-structured data:

- multitype branching processes relying on local dependency properties for analyzing motifs.
- multiple change-point models relying on long-term dependencies for segmenting trees in homogeneous zones.

In multitype branching processes, the plant development is viewed as a demographic process, a parent entity of a given type generating child entities of different types (e.g. vegetative and flowering entities). Formally, the botanical entity properties are summarized as a categorical state variable. The number of child entities in each state is modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious generation distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on exploring the space of possible chain graph models, which required defining a notion of neighbourhood of these graphs [59]. To relax the strong constraints regarding dependencies induced by parametric distributions, mixture of graphical models were also considered [60]. Multitype branching processes were applied to the analysis of the patchiness pattern (consisting of canopies made of clumps of either vegetative or flowering growth units) in mango trees. To identify the clumps, a novel approach based on tree-segmentation was developed [35].

- *Simulating fruit tree phenology* (A.S. Briand, Frédéric Boudon, Frédéric Normand [CIRAD, HortSys, Réunion Island], Anaëlle Dambreville, Jean-Baptiste Durand, Pierre Fernique, Yann Guédon, Christophe Pradal, Pierre-Eric Lauri [AFEF team, AGAP])

Mango is a tropical tree characterized by strong asynchronisms within and between trees. To study more precisely the interplay between the plant structural components, we built an integrative model to simulate the plant development based on the L-system formalism and GLM to model the dependencies between events. With such model, we showed the importance of architectural and temporal factors in the development of the units of the trees. The model also simulates the phenology of shoots and inflorescences. For this, the sizes of the different organs is modelled by statistical laws estimated from measurements that depends on their locations in the architecture. The growth speed of organs is modulated by the temperature. This structural and phenological model has been presented at the ISHS symposium on Montpellier [37].

This year, the model has been extended, during the intership of S. Persello to take into account fruiting probabilities and coupled with an ecophysiological model of fruit growth [68], [69]. The global aim is to have a crop simulation model to predict fruit yield and quality on mango tree. An overview of this global model based on the coupling of different structural or ecophysiological sub-models has been also presented in different ISHS symposia [40], [50]

- *Integrative developmental growth stages of shoots* (Anaëlle Dambreville, Yann Guédon, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand [CIRAD, HortSys, Réunion Island])

Plant growth, i.e. the increase of organ dimensions over time, and development, i.e. the change in plant structure, are often studied as two separate processes. However, there is structural and functional evidence that these two processes are strongly related. Our aim was to investigate the co-ordination between growth and development using mango trees, which have well-defined developmental stages. Developmental stages, determined in an expert way, and organ sizes, determined from objective measurements, were collected during the vegetative growth and flowering phases of two cultivars of mango. For a given cultivar and growth unit type (either vegetative or flowering), a multi-stage model based on absolute growth rate (AGR) sequences deduced from the measurements was first built, and then growth stages deduced from the model were compared with developmental stages. Strong matches were obtained between growth stages and developmental stages, leading to a consistent definition of integrative developmental growth stages [14]. The growth stages highlighted growth asynchronisms between two topologically connected organs, namely the vegetative axis and its leaves. Integrative developmental growth stages emphasize that developmental stages are closely

related to organ growth rates and can be interpreted in terms of the possible physiological processes (hydraulics, biomechanics and carbohydrate partitioning) underlying these stages. We also explore growth stages deduced from relative growth rate (RGR) sequences applying the same methodology. AGR and RGR have different meanings regarding plant metabolism since AGR represents net sink strength whereas RGR represents net sink activity. For vegetative growth units, the match rates between RGR-based stages and developmental stages were rather similar to the match rates between AGR-based stages and developmental stages, because of the rich information provided by the four organs modeled (the axis and three selected leaves). The match rates were far lower for the inflorescences where only the main axis was modeled. This is related to the fact that, compared to AGRs, RGRs amplify the variations at the beginning of growth of an organ while damping the variations at the end of growth.

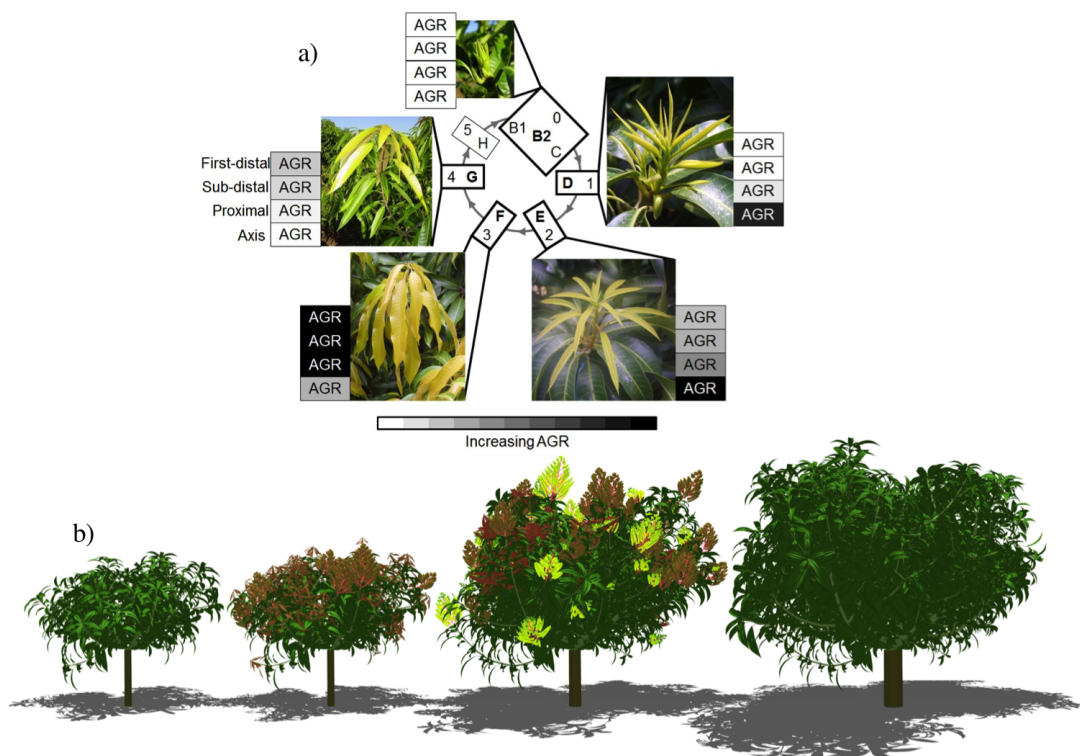


Figure 1. a) Correspondences between developmental stages of mango growth units determined from morphological observations in an expert way (inside the central circle) and growth stages obtained using segmentation models (outside the circle) [14]. Main developmental stages in bold are illustrated by the photographs. Variations in absolute growth rate for axis and leaves are illustrated using a white to black scale. b) Simulation of the development of a mango tree over two cycles [37]. The first and last image corresponds to the end of the vegetative period of the 3rd and 5th growing cycle (June), respectively while the second and third images correspond to the flowering phase (August) of the 3rd and 4th cycles, respectively. The different colours of the inflorescences of the 3rd image show different developmental stages and the flowering asynchronism over the tree.

- Characterizing the successive flowering phases of strawberry in relation to genetic determinants (Yann Guédon, Marc Labadie, Béatrice Denoyes [INRA, UMR BFP, Villenave d'Ornon], Justine Perrotte)

Our aim was to characterize the successive flowering phases of perpetual flowering strawberry genotypes, which is of particular importance for better predicting fruit production. We applied multiple change-point models for the synchronous segmentation of the individuals of a given genotype in successive flowering phases. We identified two groups of genotypes that differ by the intensity of the flowering at the end of the flowering period. Using a genetic approach, we identified a locus controlling the flowering intensity at the end of the flowering period that likely explain these two groups of genotypes. A multivariate generalization of the synchronous segmentation approach is developed in the context of Marc Labadie's PhD, the idea being to characterize not only the flowering pattern as in our first study but more generally the developmental pattern combining vegetative development, branching and flowering.

- *Self-nested structure of plants.* (Christophe Godin, Romain Azaïs, Farah Ben Naoum, Jean-Baptiste Durand, Alain Jean-Marie)

In a previous work [7], we designed a method to compress tree structures and to quantify their degree of self-nestedness. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG (Directed Acyclic Graph, equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree.

The method proposed in [7] thus compresses a tree in width, but not in height. In a new work, we designed an extension of this compression method in which a tree is compressed in both width and height. The method is based on the detection of so-called *quasi-isomorphic paths* in a tree and on the compression of these paths in height. A paper describing the corresponding algorithms has been recently accepted in the Journal of Theoretical Biology (To appear).

6.1.3. Analyzing the influence of the environment on the plant ontogenic programme

Participants: Jean-Baptiste Durand, Christian Fournier, Christophe Godin, Yann Guédon, Christophe Pradal, Jean Peyhardi, Pierre Fernique, Guillaume Garin.

This research theme is supported by three PhD programs.

The ontogenetic programme of a plant is actually sensitive to environmental changes. If, in particular cases, we can make the assumption that the environment is a fixed control variable (see section 6.1.2), in general the structure produced by meristem results from a tight interaction between the plant and its environment, throughout its lifetime. Based on observations, we thus aim to trace back to the different components of the growth (ontogenetic development and its modulation by the environment). This is made using two types of approaches. On the one hand, we develop a statistical approach in which stochastic models are augmented with additional time-varying explanatory variables that represent the environment variations. The design of estimation procedures for these models make it possible to separate the plant ontogenetic programme from its modulation by the environment. On the other hand, we build reactive models that make it possible to simulate in a mechanistic way the interaction between the plant development and its environment.

- *Influence of environmental conditions and horticultural practices on the branching and axillary flowering structures of fruit tree shoots.* (Yann Guédon, Evelyne Costes [AFEF Team, AGAP], Ted DeJong [UC Davis], Claudia Negron [UC Davis]).

In the context of a collaboration with Claudia Negron and Ted DeJong, we studied the influence of water availability and pruning practices [21] on the branching and axillary flowering structures of different categories of almond shoots. Stochastic models (hidden semi-Markov chains) were built for the branching and axillary flowering structures of different categories of almond shoots corresponding to different genetic backgrounds, levels of irrigation and pruning practices.

- *Analyzing growth components in trees.* (Yann Guédon, Yves Caraglio [AMAP], Olivier Taugourdeau [AMAP])

We identified robust indicators that summarize the respective importance of ontogeny and environmental constraints (mainly related to light environment) in forest tree development [26]. In this context, tree growth data correspond to the retrospective measurement of annual shoot characteristics (e.g. length, number of branches) along the main stem. We applied segmentation models to identify tree growth phases. These segmentation models, which are hidden semi-Markov chains, were compared with simple hidden Markov chains that correspond to the environment-driven development assumption. This statistical modelling approach was applied to both evergreen (Corsican pine and silver fir) and deciduous (sessile oak and Persian walnut) tree species growing in contrasted conditions ranging from managed forest stands to unmanaged understoreys. Growth phase duration distributions estimated within these segmentation models characterize the respective importance of ontogeny and environmental constraints in tree development at the population scale and have very contrasted characteristics in terms of shape and relative dispersion between ontogeny-driven and environment-driven tree development. These characteristics may change over tree life, reflecting changes in tree competition. Growth phase duration distributions summarize the joint trajectory of tree ontogeny and environment without requiring tree growth follow-up data for their estimation.

- *Analyzing fruit tree phenology in various climatic conditions* Yann Guédon, Jean-Michel Legave [AFEF team, AGAP], Gustavo Malagui [Universidade Tecnológica Federal do Paraná]

The responses of flowering phenology to temperature increases in temperate fruit trees have rarely been investigated in contrasting climatic regions. This is an appropriate framework for highlighting varying responses to diverse warming contexts, which would potentially combine chill accumulation declines and heat accumulation increases. To examine this issue, a data set was constituted in apple tree from flowering dates collected for two phenological stages of three cultivars in seven climate-contrasting temperate regions of Western Europe and in three mild regions, one in Northern Morocco and two in Southern Brazil. Multiple change-point models were applied to flowering date series, as well as to corresponding series of mean temperature during two successive periods, respectively determining for the fulfillment of chill and heat requirements. A new overview in space and time of flowering date changes was provided in apple tree highlighting not only flowering date advances as in previous studies but also stationary flowering date series [18]. At global scale, differentiated flowering time patterns result from varying interactions between contrasting thermal determinisms of flowering dates and contrasting warming contexts. This may explain flowering date advances in most of European regions and in Morocco vs. stationary flowering date series in the Brazilian regions. A notable exception in Europe was found in the French Mediterranean region where the flowering date series was stationary. While the flowering duration series were stationary whatever the region, the flowering durations were far longer in mild regions compared to temperate regions. Our findings suggest a new warming vulnerability in temperate Mediterranean regions, which could shift towards responding more to chill decline and consequently experience late and extended flowering under future warming scenarios.

- *Investigating how architectural development interfere with epidemics and epidemic control* (Christian Fournier, Corinne Robert [Ecosys, INRA], Guillaume Garin [ITK, Montpellier], Bruno Andrieu [Ecosys, INRA], Christophe Pradal)

Sustainable agriculture requires the identification of new, environmentally responsible strategies of crop protection. Modelling of pathosystems can allow a better understanding of the major interactions inside these dynamic systems and lead to innovative protection strategies. In particular, functional–structural plant models (FSPMs) have been identified as a means to optimize the use of architecture-related traits. A current limitation lies in the inherent complexity of this type of modelling, and thus the purpose of this work is to provide a framework to both extend and simplify the modelling of pathosystems using FSPMs. Complex models are disassembled into separate *knowledge sources* originating from different specialist areas of expertise and these can be shared and reassembled into multidisciplinary models. This year, we worked on four application studies that used the framework. In the frame of the PhD of Guillaume Garin, we perform a validation of the wheat septoria model, an analysis of the influence of the wheat architecture on the competition

between septoria and brown rust, and a sensitivity analysis of the response of the severity of septoria to architectural traits. In the frame of the Echap project, we use the wheat-septoria model to indentify optimal date of pesticide application. All these studies allows to populate the framework with consistent example of application, and lead to the development of operational modules that allows the fitting and validation of pathosystem models with experimental data.

6.2. Meristem functioning and development

In axis 2 work focuses on the creation of a *virtual meristem*, at cell resolution, able to integrate the recent results in developmental biology and to simulate the feedback loops between physiology and growth. The approach is subdivided into several sub-areas of research.

6.2.1. Data acquisition and design of meristem models

- *Improvement of the MARS-ALT pipeline robustness* Meristem, laser microscopy, image reconstruction, cell segmentation, automatic lineaging

Participants: Léo Guignard, Christophe Godin, Christophe Pradal, Grégoire Malandain [Morpheme, Inria], Gaël Michelin [Morpheme, IPL Morphogenetics, Inria], Guillaume Baty, Sophie Ribes [IBC, UM], Jan Traas [RDP, ENS], Patrick Lemaire [CRBM, CNRS], Yassin Refahi [RDP, ENS-Lyon / Sainsbury Lab, Cambridge, UK].

This research theme is supported by a PhD FRM grant, Jan Traas's ERC, Inria ADT programme and the Morphogenetics Inria Project Lab.

The MARS-ALT (Multi-Angles Registration and Segmentation - Automatic Lineage Tracking) software pipeline [6] automatically performs a segmentation at cell resolution from 3D or 2D voxel images where the membranes/walls are marked (by a die for example) and makes it possible to follow the lineage of these cells through time.

This year, the ALT tracking pipeline has been reformulated by using a generic cell modeling approach (enabling for example more than one cell division), and both stability and robustness were improved. The modeling approach is generic and can be used on other kind of data (nuclei, human cells, ...). These trials will be conducted during the year. Moreover, the architecture of the image processing components has been modified (plugin approach) and integrated with the TissueLab platform. Some vizualisation tools have been improved, and the platform includes a module allowing an interaction with data (Alizon Konig, master internship). This point enables an efficient creation of gold standard to validate segmentation results.

This year, we also finalize the development of a new segmentation and tracking pipeline, ASTEC (Adaptive Segmentation and Tracking of Embryonic Cells). ASTEC is a one-pass algorithm (in contrast to MARS-ALT, that perform first the segmentation and then the tracking in two-passes) that is best suited for movies with numerous close time-points acquired at high spatio-temporal resolution. This pipeline takes advantage of information redundancy across the movies and biological knowledge on the segmented organism to constrain and improve the segmentation and the tracking. We used this one-pass algorithm to segment and track all cell shapes of a developing embryo of the marine invertebrate *Phallusia mammillata*. As a result we obtained the full track of the shapes of all the cells from the 64 cell stage up to the early tailbud stage (1030 cells undergoing 640 division events followed across 180 time-points through 6 hours of development imaged every 2 minutes, Figure 2).

Based on this quantitative digital representation, we systematically identified cell fate specification events up to the late gastrula stage. Computational simulations revealed that remarkably simple rules integrating measured cell-cell contact areas with spatio-temporal expression data for extracellular signalling molecules are sufficient to explain most early cell inductions. This work suggests that in embryos developing with stereotyped cell shapes and positions (like *Phallusia mammillata* embryos), the genomic constraints for precise gene expression levels are relaxed, thereby allowing rapid genome evolution.

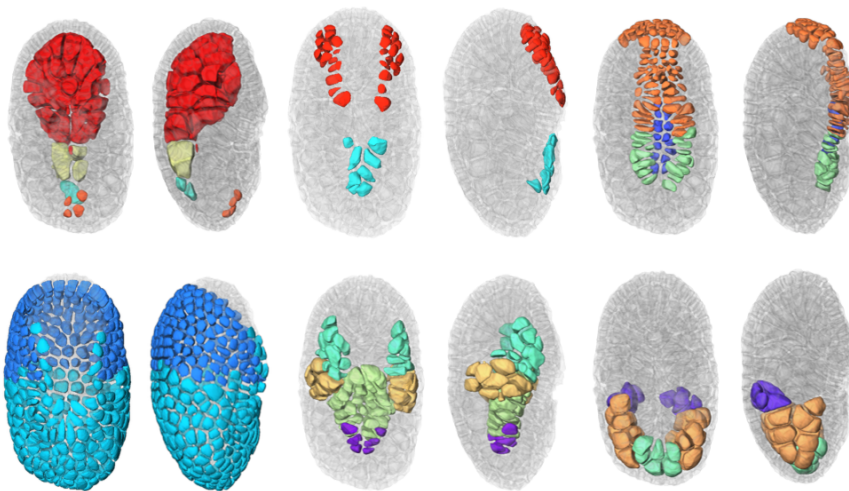


Figure 2. 3D projection of the segmented embryo at the early tailbud stage. The cells are colored by tissue type. The cells are slightly eroded to allow their distinction. The other cells of the embryo are in transparent grey. The dorsal and lateral sides are shown.

- *Creating mesh representation of cellular structures*

Participants: Guillaume Cerutti, Sophie Ribes, Christophe Godin, Géraldine Brunoud [RDP, ENS], Carlos Galvan-Ampudia [RDP, ENS], Teva Vernoux [RDP, ENS], Yassin Refahi [RDP, ENS, Sainsbury Lab].

This research theme is supported the HFSP project Biosensors.

To produce a more efficient data structure accounting for the geometry of cellular tissues, we studied the problem of reconstructing a mesh representation of cells in a complex, multi-layered tissue structure, based either on membrane/wall images segmented using MARS or on nuclei images of shoot apical meristems. The construction of such mesh structures for plant tissues is currently a missing step in the existing image analysis pipelines.

We developed tools to reconstruct a 3D cell complex representing the tissue, based on the dual simplicial complex of cell adjacencies. This set of tetrahedra is optimized from a reasonable initial guess to match the adjacencies in the tissue, which proved to produce a very faithful reconstruction [39]. We also developed a set of methods to triangulate such reconstructions, and enhance the quality of triangular mesh representations of plant tissue, simultaneously along several criteria [28].

These tools can produce light discrete representations of the cell interfaces that enables fast visualization, information projection, and quantitative analysis of the tissue, and have given way to some of the first biomechanical simulations on real-world data.

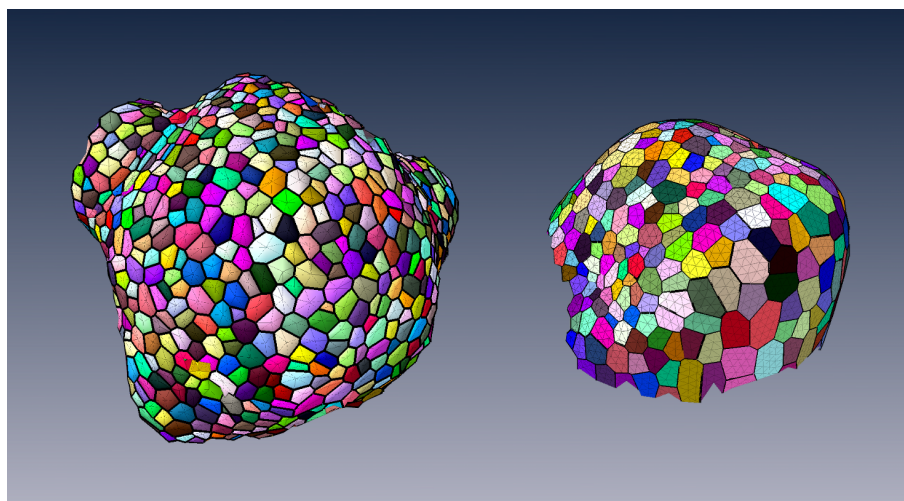


Figure 3. Triangular mesh representations of shoot apical meristem and flower meristem tissues obtained from MARS segmentations

- *Design of 3D digital atlases of tissue development*

Participants: Sophie Ribes, Yassin Refahi [RDP, ENS, Sainsbury Lab], Guillaume Cerutti, Christophe Godin, Christophe Pradal, Christophe Pradal, Frédéric Boudon, Gregoire Malandain [RDP, ENS], Gaël Michelin [RDP, ENS], Guillaume Baty, Jan Traas [RDP, ENS], Teva Vernoux [RDP, ENS], Patrick Lemaire [CRBM, CNRS], Françoise Monéger [RDP, ENS].

This research theme is supported the Inria Project Lab Morphogenetics, the ADT Mars-Alt and the HFSP project Biosensors.

To organize the various genetic, physiological, physical, temporal and positional informations, we build a spatialized and dynamic database [67]. This database makes it possible to store all

the collected information on a virtual 3D structure representing a typical organ. Each piece of information has to be located spatially and temporally in the database. Tools to visually retrieve and manipulate the information, quantitatively through space and time are being developed. For this, the 3D structure of a typical organ has been created at the different stages of development of the flower bud. This virtual structure contains spatial and temporal information on mean cell numbers, cell size, cell lineages, possible cell polarization (transporters, microtubules), and gene expression patterns. Such 3D digital atlas is mainly descriptive. However, like for classical databases, specific tools make it possible to explore the digital atlas according to main index keys, in particular spatial and temporal keys. Both a dedicated language and a 3D user interface are being designed to investigate and query the 3D virtual atlas. Current developments of this tool consist in using directly the segmented images produced from laser microscopy to build the atlas. To better represent the development of a biological population, a method to compute an "average" structure is investigated.

6.2.2. Shape analysis of meristems

Participants: Jonathan Legrand, Pierre Fernique, Frédéric Boudon, Yann Guédon, Christophe Godin, Pradeep Das [RDP, ENS], Arezki Boudaoud [RDP, ENS].

At cellular resolution, we studied the organization of cells in the meristems. The MARS-ALT pipeline provides rich spatio-temporal data sets for analyzing the development of meristems. A first step consisted of designing a dedicated graph structure for efficiently representing the spatial (adjacency between cells) and temporal (cell division) relationships between cells. Various variables can be attached either to the vertices (e.g. cell volume, inertia axes) or the edges (e.g. wall surface, distance between cell centroids). This graph may be augmented by new variables resulting from various spatial or temporal filtering (e.g. cell volumetric growth). Looking at homogeneous regions in the variable value space, cellular patterns can be identified.

Considering the highly-structured nature of our data (time and space structuring) and the potential diversity and heterogeneity of possible cell descriptors, we developed two complementary approaches:

- A first one that favours the spatial structuring: In this approach, the cell neighbourhood and the cell descriptors are jointly taken into account in a clustering approach whose objective is to identify a small number of clusters corresponding to well-defined cell identities. Once the cells have been labelled using the clustering algorithm, cell generation distributions are estimated on the basis of the labelled lineage trees.
- A second one that favours the temporal structuring: In this approach, the data of interest are lineage forest and the only spatial structuring taken into account corresponds to siblings with respect to a given parent cell. In a first step, cell identities are inferred on the basis of the cell descriptors taking into account lineage relationships using hidden Markov tree models and the spatial regions that emerge from the cell identity labelling are then characterized. This second approach is supported by the fact that cell topology is only affected by division which makes highly relevant the local spatial information taken into account in this approach.

6.2.3. Mechanical models of plant tissues

Participants: Jean-Philippe Bernard, Olivier Ali, Christophe Godin, Benjamin Gilles, Frédéric Boudon, Ibrahim Cheddadi, Jan Traas [ENS-Lyon], Olivier Hamant [ENS-Lyon], Arezki Boudaoud [ENS-Lyon].

This research theme is supported by the Inria Project Lab Morphogenetics and the Jan Traas's ERC.

The rigid cell walls that surround plant cells are the main load-bearing structures in plant tissues. These walls are submitted to stresses due to cell turgor pressure. Above some threshold, these stresses cause deformation in the cell walls and triggers wall **irreversible expansion** (*synthesis*). Shape changes of plant tissues are therefore tightly related to the turgidity of cells and to the mechanical state and the molecular composition of the underlying cell walls. We developed a conceptual and numerical framework to model the mechanical structure of cell walls and their deformation by turgor pressure in 3-dimensions. This framework was used to study the interplay between post-transcriptional regulation, biochemistry, and mechanics within growing plant tissues. This work has been published this year in Plos Computational Biology [13].

In this first step, all mechanical and structural quantities are defined at the tissular scale. This is made possible by abstracting the connection between the actual molecular composition of the walls and the various signalling cascade at play during growth. To extend this approach, we also started to develop a mechanobiological approach relating the irreversible expansion of the walls to molecular mechanisms happening within them, based on the thermodynamical equilibrium of the pectin-based matrix within the wall. We propose that at the molecular scale expansion of this matrix is based on the adsorption of newly synthesized pectin molecules. This adsorption mechanism is regulated by the mechanical stresses applied on the wall. We show that this mechanism belongs to a class of biochemical / biomechanical processes commonly appearing in the dynamics of supra-molecular load-bearing structures: the force-driven polymerization processes. A preliminary version of these ideas (the 1D case) is currently under review in *Trends In Plants Sciences*.

We also considered to extend the original modeling approach to situations where entire organ dynamics should be modeled over large time lapse (several days) (PhD work of Jean-Philippe Bernard). In our first approach, the mechanical model relies on a finite element method (FEM) to describe the deformation of the tissue. In FEM, the tissue is represented by a mesh. The positions of the vertices at each time step are estimated from a linear system. If the tissue is big or if the mesh is fine, the linear system can be large and thus leads to computational overheads. An alternative way to classical FEM is to use a meshless method where the deformation of the tissue can be characterized by a linear combination of deformations of a finite and small set of frames. Because shape functions are no longer defined on each element but on the whole tissue, they have to be updated at each growth step by estimating a new rest configuration. With meshless method, the discretization of the system can be dynamically updated parsimoniously according to the precision required to model the emergence of shapes. With an uniform distribution of the frames within the volume, our method still leads to computational overheads. However, since the meristem initiates a branching structure at a macroscopic scale, we combined our mechanical model at tissular resolution with classical method used to generate branching structures at macroscopic scales. For this, we use the information of the plant branching structure to distribute the frames along the plant's axes. This allows us to use curvilinear shape functions while describing the branching structure growth using L-systems. This multi-scale framework allows us to define developmental rules which can initiate new organs at the surface of the meristematic dome by softening locally the meristem dome and thus creating new growing initia. First very encouraging results were obtained this year that demonstrate the feasibility of the approach.

6.2.4. Gene regulatory networks: Design of a genetic model of inflorescence development.

Participants: Eugenio Azpeitia, Christophe Godin, François Parcy, Etienne Farcot.

This research theme is supported by the Inria Project Lab Morphogenetics.

Modeling gene activities within cells is of primary importance since cell identities correspond to stable combination of gene expression.

We studied the regulatory network that controls the flowering transition during morphogenesis. To overcome the network complexity and integrate this regulation during ontogenesis, we have developed a first model of the control of floral initiation by genes, and in particular the situation of cauliflower mutants, in which the meristem repeatedly fails in making a complete transition to the flower. Three different network models were done and validated. A first Boolean version, a second fuzzy logic and an ODEs models were studied. The models are able to correctly recover the gene steady states observed in the meristems during the flower transitions, the gene transitions and the mutant effects. Importantly, the model is able to explain the cauliflower mutants. This work couples models at different scales, since the gene regulatory network is used as a decision module in an L-system model of the inflorescence architecture. This mixed model has led us to make different hypotheses about gene interactions and hormonal regulation. First predictions about gene actors controlling the passage to flower could be verified. Some links between gene regulation and plant growth have been identified. These links can be experimentally tested which could lead to a first integrated picture of flower development.

Finally, given that the cauliflower have different morphologies (i.e. regular and romanesco cauliflower morphologies) we explored the effect of changes in the L-system parameter values over the cauliflower morphology. Interestingly, we discovered by exploring the model that variations in the regulation of some phyllotactic

parameters can produce the different cauliflower morphologies and explain other reported differences among them. Predictions were made using the model and experimental validations of this hypothesis are currently being tested. All our results could provide a comprehensive understanding of how genes and plant architecture are linked in a dynamical way.

6.2.5. *Modelling the influence of dimerisation sequence dissimilarities on the auxin signalling network*

Participants: Jonathan Legrand, Yann Guédon, Teva Vernoux [ENS-Lyon].

Auxin is a major phytohormone involved in many developmental processes by controlling gene expression through a network of transcriptional regulators. In *Arabidopsis thaliana*, the auxin signalling network is made of 52 potentially interacting transcriptional regulators, activating or repressing gene expression. All the possible interactions were tested in two-way yeast-2-hybrid experiments. Our objective was to characterise this auxin signalling network and to quantify the influence of the dimerisation sequence dissimilarities on the interaction between transcriptional regulators. We applied model-based graph clustering methods relying on connectivity profiles between transcriptional regulators. Incorporating dimerisation sequence dissimilarities as explanatory variables, we modelled their influence on the auxin network topology using a mixture of linear models for random graphs. Our results provide evidence that the network can be simplified into four groups, three of them being closely related to biological groups. We found that these groups behave differently, depending on their dimerisation sequence dissimilarities, and that the two dimerisation sub-domains might play different roles. We proposed the first pipeline of statistical methods combining yeast-2-hybrid data and protein sequence dissimilarities for analyzing protein-protein interactions. We unveiled using this pipeline of analysis the transcriptional regulator interaction modes.

6.2.6. *Model integration*

Participants: Frédéric Boudon, Christophe Godin, Guillaume Baty, Guillaume Cerutti, Jean-Louis Dinh, Jan Traas.

This research theme is supported by the Morphogenetics Inria Project Lab.

Our approach consists of building a programmable tissue which is able to accept different modeling components. This includes a central data structure representing the tissue in either 2-D or 3-D, which is able to grow in time, models of gene activity and regulation, models of signal exchange (physical and chemical) between cells and models of cell cycle (which includes cell division). An introduction to the modeling of some main components of such integrated system was published as a book chapter in the series of *Ecole de Physique des Houches* [43]. For each modeling component, one or several approaches are investigated in depth, possibly at different temporal and spatial scales, using the data available from the partners (imaging, gene networks, and expression patterns). Approaches are compared and assessed on the same data. The objective of each sub-model component will be to provide plugin components, corresponding to simplified versions of their models if necessary, that can be injected in the programmable tissue platform. This work is developed in collaboration with the RDP group at ENS-Lyon [70] and the CPIB group in Nottingham, UK [53].

One key aspect of our approach is the development of a computer platform dedicated to programming virtual tissue development, TissueLab. This platform, based on *OpenAlea*, will be used to carry out integration of the different models developed in this research axis. In the past year, progress has been made in defining a generic tissue data structure that could be used in this platform. Currently, robust geometric operations such as division are implemented and tested. Moreover, a redesign of the structure based on more elaborated formalisms such as combinatorial maps is being investigated. A 2D version is being developed in the context of Jean-Louis's Dinh PhD thesis, and will be described in a forthcoming book chapter.

6.3. Multi-scale models and analysis: from cells to plant architecture (and back)

6.3.1. *Modeling water transport in roots*

Participants: Mikaël Lucas [IRD], Christophe Pradal, Christophe Godin, Yann Boursiac Bpmp, Christophe Maurel [BPMP].

This research theme is supported by the ANR project HydroRoot.

A model of *Arabidopsis thaliana* root hydraulics at the cellular level was developed in the OpenAlea modeling platform. The model relies on the integration throughout root architecture of elementary hydraulic components. Each component integrates local radial and axial water flows. Axial hydraulic conductivity is calculated according to Poiseuille's law, based on local size of xylem vessels. Radial hydraulic conductivity is determined in part by aquaporin activity and was set constant throughout root architecture in the first model versions. In its current state, the model is parameterized using architectural, tissular and physiological data that were experimentally determined in the Aquaporin group at UMR BPMP. The architectural reconstruction of the root system is based on a tridimensional multi-scale tree graph (MTG). The current model is capable of predicting the water flow that is transported by a root system in the standard experimental conditions used in the Aquaporin group. This model was used to perform sensitivity analyses and determine the respective contributions to root hydraulic dynamics of various biological parameters (axial and radial hydraulic conductivities, root architecture). One major finding is that the root hydraulic conductivity (L_{pr}) computed from the model is highly dependent on root architecture. This is due to the limiting role of axial (xylem) conductance, one feature that had been neglected in previous representations of root water transport. The radial hydraulic conductivity may primarily be limiting in conditions of L_{pr} inhibition, since its increase from values in control roots has marginal effects on L_{pr} . A new set of experimental data including root diameter repartitions in wild-type plants, and xylem vessel diameters in mutants with altered xylem morphology (*irx3*, *esk1*) will be used to challenge the model. Root cell hydraulic conductivities will also be measured in these and aquaporin mutant phenotypes. Our aim is to check whether, based on anatomical and morphological data, the model can properly predict the radial hydraulic conductivity of these genotypes.

As the simulations may be time consuming and results sometimes difficult to interpret on complex branching systems, we started to investigate new methods to compute efficiently hydraulic conductivities and corresponding flows on complex root systems using architecture compression techniques developed in the 1st axis of the project. First results show that very efficient computations of complex hydraulic architectures can be derived from the use of these compression techniques on idealized root architectures. These encouraging results provide a new abstraction that will be used in combination with the detailed modeling approach described above to break down the complexity of the analysis these huge branching systems.

6.3.2. Mechanical modeling of fruit growth

Participants: Ibrahim Cheddadi [Inra, Avignon], Mik Cieslak [U. Calgary], Frédéric Boudon, Valentina Baldazzi [Inra, Avignon], Nadia Bertin [Inra, Avignon], Michel Genard [Inra, Avignon], Christophe Godin.

This research theme is supported by the Agropolis project MecaFruit3D.

Fruits and plants in general are large scale hydraulic systems in which growth is closely linked to water fluxes: thanks to osmotic pressure difference, the cells are able to absorb water from their environment and therefore increase their volume; as the cells are bounded by rigid walls, this results in both hydrostatic pressure (the so-called turgor pressure) in the cell and tension in the cell walls; above a threshold, synthesis of new cell wall material occurs and relaxes the tension. This process allows cells to grow, and along with cell division, is responsible for plant growth. In fruits, phloem and xylem vascular networks provide the water fluxes necessary for growth, while the osmotic pressure is mainly regulated by sugar intake from the phloem. The goal of this project is to combine a description of water and sugar fluxes at the fruit scale (see section 4) with a modelling of growth at cell level, as described above.

As a first step in this direction, we have developed a bidimensional multicellular model that couples, on the one hand, water fluxes between cells (symplastic pathway) and between cells and intercellular space (apoplastic pathway), and on the other hand, mechanical properties of the cell walls and mechanical equilibrium of this complex system. Existing multicellular models for plant growth overlook this coupling. From a mathematical point of view, it corresponds to a coupling between (1) the ordinary differential equations that describe fluxes and cell walls properties and (2) the highly non linear system of equations that describes the mechanical equilibrium of the cell walls.

We have developed a numerical method for this coupled system, that allows to simulate in a reasonable amount of time a hundred of connected cells. The non linear system of equations (2) is the bottleneck to reach a higher number of cells; in order to overcome this, we plan to use the framework developed for the mechanical modelling of meristems (see section 6.2.3) and adapt it to this system. This will also allow to address tridimensional tissues.

Numerical simulations exhibit a highly non linear behaviour with respect to the governing parameters. We have identified two clearly distinct growth regimes: one regime that allows large growth heterogeneities by amplifying the effect of differences between cells, and conversely another regime that smoothes differences out and yields a homogeneous growth. On the biological level, the first regime is well adapted to morphogenesis, whereas the second one is well adapted to homothetic growth after the differentiated tissues have been created. A publication of these completely new results is in preparation.

We plan to compare this model to experimental results of the tomato fruit at the tissue level. In the longer term, a continuous version of this multicellular model could be an interesting way to build a model at the fruit scale.

6.3.3. Analyzing root growth and branching

Participants: Beatriz Moreno Ortega, Sixtine Passot, Yann Guédon, Laurent Laplaze [IRD, DIADE], Mikaël Lucas [IRD, DIADE], Bertrand Muller [INRA, LEPSE].

This research theme is supported by two PhD programmes.

New 2D and 3D root phenotyping platforms are emerging with associated image analysis toolbox (e.g. Smart-Root, RhizoScan) and the high-level analysis these complex phenotyping data requires new computational investigation methods.

Here, we aim at developing a pipeline of methods for analyzing root systems at three scales:

1. tissular scale to identify and characterize the division, elongation and mature zones along a root using piecewise heteroscedastic linear models. To this end, we introduced a new slope heuristic for the selection of the number of zones in cell length series [29] [36].
2. individual root scale to analyze the dynamics of lateral root elongation. We investigated the use of semi-Markov switching linear models for classifying roots on the basis of the identification of phases within growth rate profiles,
3. root system scale to analyze the branching structure.

This pipeline of analysis methods will be applied to different species (maize, millet and *arabidopsis*) and for different biological objectives (study of genetic diversity for millet and of metabolic and hormonal controls of morphogenesis for maize).

6.3.4. Analyzing shoot and leaf elongation

Participants: Maryline Lièvre, Yann Guédon, Leo Guignard, Christine Granier [INRA, LEPSE].

This research theme is supported by one PhD programme and the labex Agro project "Integrated model of plant organ growth".

This study is based on the observation that there is a lack of methods enabling the integrated analysis of the processes controlling the vegetative development in *Arabidopsis thaliana*.

The changes in leaf size and shape during ontogeny associated with the heteroblastic development is a composite trait for which extensive spatio-

temporal data can be acquired using phenotyping platforms such as PHENOPSIS. However, only part of the information contained in such data is exploited and developmental phases are usually defined using a selected organ trait. We introduced new methods for identifying developmental phases in *Arabidopsis* rosette using various traits and minimum a priori assumptions. A first pipeline of analysis was developed, combining image analysis and statistical models to integrate morphological, shape, dimensional and expansion dynamics traits for the successive leaves of the *Arabidopsis* rosette. Dedicated segmentation models called semi-Markov switching models were built for selected genotypes in order to identify rosette developmental phases. Four

successive developmental phases referred to as seedling, juvenile, transition and adult were identified for the different genotypes. We showed that the degree of covering of the leaf abaxial surface with trichomes is not sufficient to define these developmental phases. Using our pipeline of analysis, we were able to identify the supplementary seedling phase and to uncover the structuring role of various leaf traits. This enabled us to compare on a more objective basis the vegetative development of *Arabidopsis* mutants.

We developed a second pipeline of analysis methods combining a semi-automatic method for segmenting leaf epidermis images based on the ilastik software, and the analysis of the obtained cell areas using a gamma or inverse Gaussian mixture models whose component parameters are tied by a scaling rule. These mixture models allowed us to estimate the distribution of the number of endocycles. We highlighted in this way that the mean number of endocycles changes drastically with leaf rank. We extended the inference approach to take into account not only complete cell areas but also censored cell areas (corresponding to cells that intercept the edges of the images). We also investigated possible temporal interpretations of endoreduplication using stochastic processes.

6.3.5. A stochastic model of phyllotaxis

Participants: Yassin Refahi, Christophe Godin, Etienne Farcot, Teva Vernoux [RDP, ENS].

This research theme has been supported by IBC and the Inria Project Lab Morphogenetics.

The geometric arrangement of lateral organs along plant stems, named phyllotaxis, shows a variety of striking patterns with remarkable regularities and symmetries. This has interested biologists, physicists, mathematicians and computer scientists for decades. These studies have led to a commonly accepted standard interpretation of phyllotaxis that postulates that organs inhibit the formation of new organs in their vicinity. At a molecular scale, these inhibitory fields have been shown to result from the spatio-temporal distribution of the plant hormone auxin. This model theoretically explains a large part of the diversity of phyllotactic patterns observed in plants.

Recently, our colleagues from ENS-Lyon observed intriguing perturbation in *arabidopsis* mutants. These perturbations were also present, to a lesser extent in the wild type. In a series of works [74], [64], [2], we could show that these perturbations patterns in both wild-type and mutant plants could be explained by permutations in the order of insertion along the stem of 2 or 3 consecutive organs. After closer inspection, we realized that the mutated gene encodes a protein diffusing from the organs and creating a field around the organs that regulates the plastochron. We could demonstrate that in the mutant, the absence of this field leads to co-initiations and subsequently to the observed permutations.

To proceed further and find a mechanistic interpretation of this phenomenon, we developed a stochastic extension of the standard model of phyllotaxis. We first analyzed the properties of the inhibitory fields created by the existing primordia on the initiation of new primordia, and concluded that the angular positions of organs are very robust to perturbations while plastochrons may be dramatically affected. This suggested that there exists a strong decoupling between space and time in the patterning process. To account for this observation, we modeled the perception of the initiation signal by cells using stochastic processes coupled with the intensity of inhibitory fields and showed that the observed permutation patterns emerge spontaneously from this purely local processes. This model recapitulates accurately the classical phyllotactic patterns and, in addition, produces realistic pattern disorders at higher organization levels as a result of stochasticity in signal perception. We show that these subtle disorders surprisingly reveal key information on the functioning of the developmental system and can therefore be regarded as *biological watermarks* of the system. In genetically or environmentally modified plants, these biological watermarks inform us on the molecular mechanisms that have been affected in the experiment. Our theoretical analysis allows us to predict the specific pattern variations that would arise from perturbations of the signaling pathways involved in lateral inhibition signaling at the shoot apex. A paper describing this model has been submitted recently for publication.

6.3.6. The role of auxin and sugar in rose bud outgrowth control

Participants: Jessica Bertheloot [INRA, Angers], Frédéric Boudon, Christophe Godin.

Auxin in the stem is known to be a key regulator of apical dominance. Over the last decades, many studies have been undertaken to understand its action mode, which is indirect because auxin in the main stem does not enter into the bud. Recently, apical dominance over basal buds in pea has been related to low sugar availability caused by high sugar demand of growing apical organs. Auxin and sugar are two signals regulating the entrance of bud into sustained growth in opposite ways. In the last year, it has also been demonstrated that sugar effect on bud outgrowth was preceded by a modification of the hormonal levels involved in bud outgrowth, which suggests that auxin and sugar pathways do interact in a non-trivial way. However, auxin and sugar effects have been studied separately until now. In this work, we investigate what is the combined effect of sugar and auxin on bud outgrowth, and how they integrate to regulate bud entrance into sustained growth. For this, a series of experiments has been carried out on a single-node cuttings of *Rosa hybrida* grown *in vitro* in which different combinations of sugar and auxin levels have been tested. A model of the regulatory networks controlling stem-bud molecular interaction is currently being developed.

6.4. Generic methodological results

In the context of our research work on biological questions, we develop concepts and tools in mathematics, statistics and computer science. This paragraph is intended to put emphasis on the most important results obtained by the team during the current year in these disciplines, independently of their biological application.

6.4.1. Scientific workflows

Participants: Christophe Pradal, Sarah Cohen-Boulakia, Christian Fournier, Didier Parigot [Inria, Zenith], Patrick Valduriez [Inria, Zenith].

6.4.1.1. OpenAlea scientific workflows

Analyzing biological data may involve very complex and interlinked steps where several tools are combined together. Scientific workflow systems have reached a level of maturity that makes them able to support the design and execution of such *in-silico* experiments, and thus making them increasingly popular in the bioinformatics community (e.g. to annotate genomes, assemble NGS data, ...). However, in some emerging application domains such as system biology, developmental biology or ecology, the need for data analysis is combined with the need to model complex multi-scale biological systems, possibly involving multiple simulation steps. This requires the scientific workflow to deal with retro-action to understand and predict the relationships between structure and function of these complex systems. In collaboration with the Zenith EPI, we have proposed a conceptualisation of OpenAlea workflows [34] by introducing the concept of higher-order dataflows as a means to uniformly combine classical data analysis with modeling and simulation. Ongoing work include deploying OpenAlea workflows on a Grid technology using the SciFloware middleware in close collaboration with Zenith within IBC and INRA Phenome projects.

6.4.1.2. Querying Scientific workflows repositories

Several workflow systems have developed scientific workflow repositories (e.g., repositories of Galaxy workflows at IBC, or repositories of OpenAlea workflows). Such repositories have grown to sizes that call for advanced methods to support workflow discovery, in particular for similarity search. Effective similarity search requires both high quality algorithms for the comparison of scientific workflows and efficient strategies for indexing, searching, and ranking of search results. Yet, the graph structure of scientific workflows poses severe challenges at each of these steps. We present a complete system for effective and efficient similarity search in scientific workflow repositories, based on the Layer Decomposition approach to scientific workflow comparison. Layer Decomposition specifically accounts for the directed dataflow underlying scientific workflows and, compared to other state-of-the-art methods, delivers best results for similarity search at comparably low runtimes. Stacking Layer Decomposition with even faster, structure-agnostic approaches allows us to use proven, off-the-shelf tools for workflow indexing to further reduce runtimes and scale similarity search to sizes of current repositories [25]. Very efficient and powerful ranking methods have been used in this work. We based our choice on the large scale study of algorithms for rank aggregation with ties we performed [56].

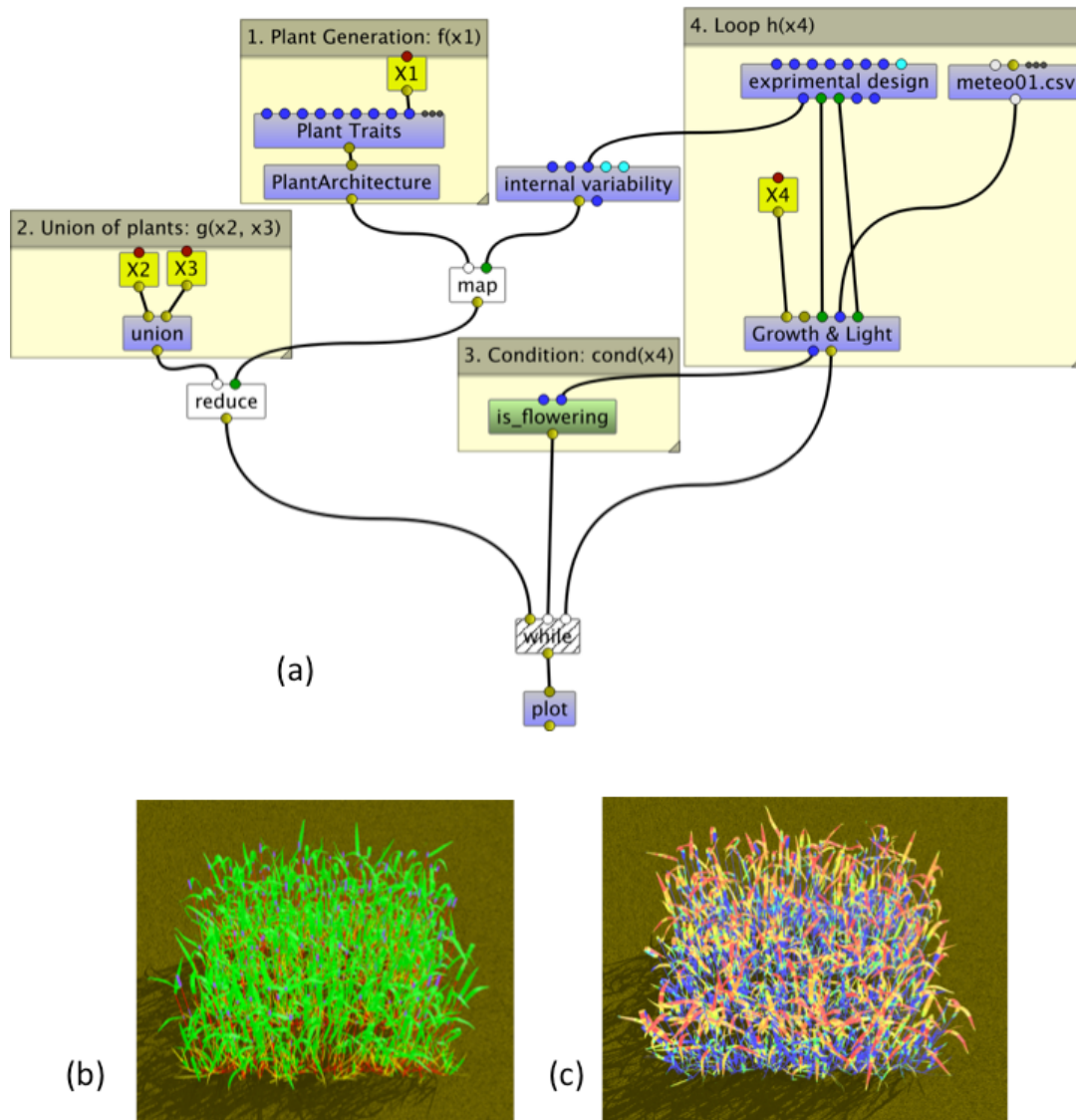


Figure 4. (a) OpenAlea workflow [34] for simulating Maize and Wheat crop performance based on phenotypic and environment data, and two image outputs (b and c). Colors represent the organ's type in (b) and the amount of intercepted light in (c).

6.4.2. *Statistical modeling*

Participants: Yann Guédon, Jean Peyhardi.

We develop statistical models and methods for identifying and characterizing developmental patterns in plant phenotyping data. Phenotyping data are very diverse ranging from the tissular to the whole plant scale but are often highly structured in space, time and scale. Problems of interest deal with the definition of new family of models specifically adapted to plant phenotyping data and the design of new methods of inference concerning both model structure, model parameters and latent structure. This is illustrated this year by [17] and [22].

6.4.3. *Lossy compression of tree structures*

Participants: Christophe Godin, Romain Azais, Jean-Baptiste Durand, Alain Jean-Marie.

the degree of self-nestedness of a tree as the edit-distance between the considered tree structure and its nearest embedded self-nested version. Indeed, finding the nearest self-nested tree of a structure without more assumptions is conjectured to be an NP-complete or NP-hard problem. We thus introduced a lossy compression method that consists in computing in polynomial time for trees with bounded outdegree the reduction of a self-nested tree that closely approximates the initial tree. This approximation relies on an indel edit distance that allows (recursive) insertion and deletion of leaf vertices only. We showed in a conference paper accepted at DCC'2016 [46] with a simulated dataset that the error rate of this lossy compression method is always better than the loss based on the nearest embedded self-nestedness tree [7] while the compression rates are equivalent. This procedure is also a keystone in our new topological clustering algorithm for trees. In addition, we obtained new theoretical results on the combinatorics of self-nested structures. The redaction of an article is currently in progress.

VISAGES Project-Team

7. New Results

7.1. Image Computing: Detection, Segmentation, Registration and Analysis

7.1.1. *Symmetric Block-Matching Registration for the Distortion Correction of Echo-Planar Images*

Participants: Renaud Hédouin, Olivier Commowick, Elise Bannier, Christian Barillot.

We introduce a new approach to correct geometric and intensity distortion of Echo Planar Images (EPI) from images acquired with opposite phase encoding directions. A new symmetric block-matching registration algorithm has been developed for this purpose relying on new adapted transformations between blocks and a symmetric optimization scheme to ensure an opposite symmetric transformation. Our results show the ability of our algorithm to robustly recover EPI distortion while obtaining sharper results than the popular TOPUP algorithm [24], [34].

7.1.2. *Quantitative analysis of T2/T2* relaxation time alteration*

Participants: Benoit Combès, Anne Kerbrat, Olivier Commowick, Christian Barillot.

T2 and T2* relaxometric data becomes a standard tool for the quantitative assessment of brain tissues and of their changes along time or after the infusion of a contrast agent. Being able to detect significant changes of T2/T2* relaxation time is an important issue. Generally, such a task is performed by comparing the variability level in the regions of interest to the variability in the normal appearance white matter. However, in the case of T2 and T2* relaxometry, this solution is highly problematic. Indeed the level of noise in the normal appearance white matter is significantly smaller than the level of noise in more intense region (e.g. MS lesions). Our aim is to provide a Bayesian analysis of T2/T2* relaxometry estimation and alteration. More specifically, we build posterior distributions for the relaxation time and the relaxation offset by elucidating the dedicated Jeffreys priors. Then the resulting posterior distributions can be evaluated using a Monte Carlo Markov Chain algorithm. Such an analysis has three main advantages over the classical estimation procedure. First it allows in a simple way to compute many estimators of the posterior including the mode, the mean, the variance and confidence intervals. Then, it allows to include prior information. Finally, because one can extract confidence interval from the posterior, testing properly whether the true relaxometry time is included within a certain range of value given a confidence level is simple.

7.1.3. *MRI quantitative imaging: Myelin Water Fraction (MWF) quantification in Multiple Sclerosis*

Participants: Olivier Commowick, Elise Bannier, Christian Barillot.

Multi-echo T2 relaxometry is potentially a relevant imaging method for MWF quantification in the study of multiple sclerosis (MS). However, to ensure accurate estimation, a large number of echoes are still required that can drive to very long acquisitions. In practice, 32 echo times ranging from 10 ms to 320 ms and an echo spacing (ESP) of 10 ms are used¹. Analysis of the decay curve of the consecutive echoes allows the estimation of the T2 spectrum. The proposed approach makes use of recent spatial regularization methods for MWF estimation from clinically compatible acquisitions (typically 11 echoes acquired within 6 minutes). The algorithms were evaluated on both synthetic and clinical data. This work was done during the internship of Lucas Soustelle [32], [29].

7.1.4. *Classification of Multiple Sclerosis Lesions using Adaptive Dictionary Learning*

Participants: Hrishikesh Deshpande, Pierre Maurel, Christian Barillot.

This work presents a sparse representation and an adaptive dictionary learning based method for automated classification of Multiple Sclerosis (MS) lesions in Magnetic Resonance (MR) images. Manual delineation of MS lesions is a time-consuming task, requiring neuroradiology experts to analyze huge volume of MR data. This, in addition to the high intra- and inter-observer variability necessitates the requirement of automated MS lesion classification methods. Among many image representation models and classification methods that can be used for such purpose, we investigate the use of sparse modeling. In the recent years, sparse representation has evolved as a tool in modeling data using a few basis elements of an over-complete dictionary and has found applications in many image processing tasks including classification. We propose a supervised classification approach by learning dictionaries specific to the lesions and individual healthy brain tissues, which include White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF). The size of the dictionaries learned for each class plays a major role in data representation but it is an even more crucial element in the case of competitive classification. Our approach adapts the size of the dictionary for each class, depending on the complexity of the underlying data. The algorithm is validated using 52 multi-sequence MR images acquired from 13 MS patients. The results demonstrate the effectiveness of our approach in MS lesion classification.

This work has been published in the journal of Computerized Medical Imaging and Graphics, Elsevier, 2015 [15]. Part of this work is published as a conference paper in ISBI 2015 [22].

7.1.5. Robust Detection of Multiple Sclerosis Lesions

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

Multiple sclerosis (MS) is a disease with heterogeneous evolution among the patients. Quantitative analysis of longitudinal Magnetic Resonance Images (MRI) provides a spatial analysis of the brain tissues which may lead to the discovery of biomarkers of disease evolution. Better understanding of the disease will lead to a better discovery of pathogenic mechanisms, allowing for patient-adapted therapeutic strategies. To characterize MS lesions, we have proposed two new approaches. The first one consists in a novel paradigm to detect white matter lesions based on a statistical framework [26]. It aims at studying the benefits of using multi-channel MRI to detect statistically significant differences between each individual MS patient and a database of control subjects. This framework consists in two components. First, intensity standardization is conducted to minimize the inter-subject intensity difference arising from variability of the acquisition process and different scanners. The intensity normalization maps parameters obtained using a robust Gaussian Mixture Model (GMM) estimation not affected by the presence of MS lesions. The second part studies the comparison of multi-channel MRI of MS patients with respect to an atlas built from the control subjects, thereby allowing us to look for differences in normal appearing white matter, in and around the lesions of each patient. Experimental results demonstrate that our technique accurately detects significant differences in lesions consequently improving the results of MS lesion detection.

Then we have presented an automatic algorithm for the detection of multiple sclerosis lesions (MSL) from multi-sequence magnetic resonance imaging (MRI) [25]. We built a probabilistic classifier that can recognize MSL as a novel class, trained only on Normal Appearing Brain Tissues (NABT). Patch based intensity information of MRI images is used to train a classifier at the voxel level. The classifier is in turn used to compute a probability characterizing the likelihood of each voxel to be a lesion. This probability is then used to identify a lesion voxel based on simple Otsu thresholding. The proposed framework was evaluated on 16 patients and our analysis reveals that our approach is well suited for MSL detection and outperforms other benchmark approaches.

7.2. Image processing on Diffusion Weighted Magnetic Resonance Imaging

7.2.1. Interpolation and Averaging of Multi-Compartment Model Images

Participants: Renaud Hédouin, Olivier Commowick, Christian Barillot.

Multi-compartment diffusion models (MCM) are increasingly used to characterize the brain white matter microstructure from diffusion MRI. We address the problem of interpolation and averaging of MCM images as a simplification problem based on spectral clustering. As a core part of the framework, we propose novel solutions for the averaging of MCM compartments. Evaluation is performed both on synthetic and clinical data, demonstrating better performance for the "covariance analytic" averaging method. We then present an MCM template of normal controls constructed using the proposed interpolation [23].

7.2.2. *The DTI Challenge: Toward Standardized Evaluation of Diffusion Tensor Imaging Tractography for Neurosurgery*

Participants: Olivier Commowick, Sylvain Prima.

Diffusion tensor imaging (DTI) tractography reconstruction of white matter pathways can help guide brain tumor resection. However, DTI tracts are complex mathematical objects and the validity of tractography-derived information in clinical settings has yet to be fully established. To address this issue, the DTI Challenge was initiated, an international working group of clinicians and scientists whose goal was to provide standardized evaluation of tractography methods for neurosurgery. The purpose of this empirical study was to evaluate different tractography techniques in the first DTI Challenge workshop. Eight international teams from leading institutions reconstructed the pyramidal tract in four neurosurgical cases presenting with a glioma near the motor cortex. Tractography methods included deterministic, probabilistic, filtered, and global approaches. Standardized evaluation of the tracts consisted in the qualitative review of the pyramidal pathways by a panel of neurosurgeons and DTI experts and the quantitative evaluation of the degree of agreement among methods. The evaluation of tractography reconstructions showed a great inter-algorithm variability. Although most methods found projections of the pyramidal tract from the medial portion of the motor strip, only a few algorithms could trace the lateral projections from the hand, face, and tongue area. In addition, the structure of disagreement among methods was similar across hemispheres despite the anatomical distortions caused by pathological tissues. The DTI Challenge provides a benchmark for the standardized evaluation of tractography methods on neurosurgical data. This study [18] suggests that there are still limitations to the clinical use of tractography for neurosurgical decision making.

7.2.3. *Diffusion MRI abnormalities detection with orientation distribution functions: A multiple sclerosis longitudinal study*

Participants: Olivier Commowick, Jean-Christophe Ferré, Gilles Edan, Christian Barillot.

We proposed a new algorithm for the voxelwise analysis of orientation distribution functions between one image and a group of reference images [13]. It relies on a generic framework for the comparison of diffusion probabilities on the sphere, sampled from the underlying models. We demonstrated that this method, combined to dimensionality reduction through a principal component analysis, allows for more robust detection of lesions on simulated data when compared to classical tensor-based analysis. We then demonstrated the efficiency of this pipeline on the longitudinal comparison of multiple sclerosis patients at an early stage of the disease: right after their first clinically isolated syndrome (CIS) and three months later. We demonstrated the predictive value of ODF-based scores for the early detection of lesions that will appear or heal.

7.3. EEG and MR Imaging

7.3.1. *On the feasibility and specificity of simultaneous EEG and ASL MRI at 3T*

Participants: Elise Banner, Marsel Mano, Isabelle Corouge, Lorraine Perronnet, Christian Barillot.

Brain functional imaging can be performed using several approaches, including EEG, BOLD and ASL MRI. To date, only a few studies have addressed the issue of connecting EEG signal to ASL perfusion. ASL imaging relies on control and label RF pulses, generating alternate gradient patterns as well as higher SAR. The aim of this study was to assess ASL-EEG at 3T in terms of safety as well as EEG and MR signal quality [19].

7.3.2. *Symmetrical EEG-fMRI Imaging by Sparse Regularization*

Participants: Pierre Maurel, Nicolas Raillard, Saman Noorzadeh, Christian Barillot.

This work [28] considers the problem of brain imaging using simultaneously recorded electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). To this end, we introduce a linear coupling model that links the electrical EEG signal to the hemodynamic response from the blood-oxygen level dependent (BOLD) signal. Both modalities are then symmetrically integrated, to achieve a high resolution in time and space while allowing some robustness against potential decoupling of the BOLD effect. The novelty of the approach consists in expressing the joint imaging problem as a linear inverse problem, which is addressed using sparse regularization. We consider several sparsity-enforcing penalties, which naturally reflect the fact that only few areas of the brain are activated at a certain time, and allow for a fast optimization through proximal algorithms. The significance of the method and the effectiveness of the algorithms are demonstrated through numerical investigations on a spherical head model. This is a joint work with T.Oberlin and R.Gribonval.

7.4. Applications in Neuroradiology and Neurological Disorders

7.4.1. Brain perfusion gender differences using ASL in young adults

Participants: Léa Itmi, Pierre Maurel, Isabelle Corouge, Jean-Christophe Ferré, Christian Barillot.

The use of population models is becoming increasingly important in cerebral imaging, particularly using Arterial Spin Labeling perfusion imaging. Therefore, it is important to know the limits of the models before applying them, to guarantee the reliability of the results. It is now well-known that brain perfusion, in particular cerebral blood flow (CBF), changes with age, and this effect needs to be taken into account when evaluating brain perfusion images. But gender differences have not been well studied yet. It is known that female brain perfusion is, in average, higher than male brain perfusion, but only few studies have investigated whether some regional perfusion differences exist or not. This work aims to assess whether, as for the age, gender differences should be taken into account when analyzing brain perfusion images. We then focus on adult subjects and study the CBF gender differences. We compared the raw CBF means and the means after normalization, we also investigated perfusion asymmetries. We used atlases for the region comparisons and the General Linear Model for the voxel level. Our results confirmed that women have a higher CBF than men, and showed that this difference can be suppressed with a normalization process, but no specific major regional difference or asymmetry was found.

7.4.2. Arterial Spin Labeling Motor Activation Presurgical Mapping for Brain Tumor Resection

Participants: Isabelle Corouge, Elise Bannier, Jean-Christophe Ferré.

Functional Arterial Spin Labeling (fASL) has demonstrated its greater specificity as a marker of neuronal activity than the reference BOLD fMRI for motor activation mapping in healthy volunteers. Motor fASL is yet to be investigated in the context of tumors, under the assumption that fASL would be less sensitive to venous contamination induced by the hemodynamics remodeling in the tumor vicinity than BOLD fMRI. As the arterial transit time may be shortened in activation areas, this preliminary study explores the ability of fASL to map the motor areas at different post-labeling delays (PLD) in healthy subjects and patient with brain tumor [21].

7.4.3. Dynamic assessment of macrophages infiltration and tissue damage in MS lesions

Participants: Anne Kerbrat, Benoit Combès, Olivier Commowick, Jean-Christophe Ferré, Elise Bannier, Christian Barillot, Gilles Edan.

Inflammation is a dynamic and complex process that could be beneficial when it supports tissue repair but also detrimental when excessive, leading to worsen tissue injury. In multiple sclerosis, it is well known from pathological and MRI studies that the prognostic between white matter lesions differed at the lesion level. Thus, 10 to 30% of T2 hyperintense lesions are seen as area of persistent hypointensity on T1-w images. These T1 hypointensity are areas of pathologically confirmed severe axonal loss. Complementary, quantitative MRI such as Diffusion imaging, magnetization transfer imaging and relaxometry can quantify and characterize

tissue changes on MRI before, during, and after the evolution of a new MRI-detected lesion. They are related to damage to myelin and axons. However, identifying in vivo the dynamic pathophysiological processes that leads to these various degree of demyelination and axonal loss in MS lesions remained challenging. In recent year, molecular and cellular imaging of the inflammatory process have been developed. Although some techniques remains at the pre-clinical level, MRI using non targeted USPIO as contrast agent can be used in MS patients. USPIO are phagocyted in periphery by macrophages and migrate to the central nervous system to characterize in vivo macrophages infiltrations within lesions. The association of cellular imaging and longitudinal quantitative MRI consist of a great opportunity to assess more specifically the overall process. In a recent study from our group, we demonstrated that infiltration of activated macrophages evidenced by USPIO enhancement, was present at the onset of MS and associated with higher local loss of tissue structure [17]. This year, we pursued this work by analyzing a longitudinal study with USPIO infusion every 3 months, associated with quantitative MRI assessment including MTI, diffusion imaging and relaxometry with the objectives of describing relationships between macrophages infiltration and quantitative MRI metrics reflecting tissue structure along time.

7.4.4. The effect of water suppression on the hepatic lipid quantification, as assessed by the LCModel, in a preclinical and clinical scenario

Participant: Elise Bannier.

This work investigates the effect of water suppression on the hepatic lipid quantification, using the LCModel. MR spectra with and without water suppression were acquired in the liver of mice at 4.7 T and patients at 3 T, and processed with the LCModel. The Cramer-Rao Lower Bound (CRLB) values of the seven lipid resonances were determined to assess the impact of water suppression on hepatic lipid quantification. A paired t test was used for comparison between the CRLBs obtained with and without water suppression. For the preclinical data, in the high (low) fat fraction subset an overall impairment in hepatic lipid quantification, i.e. an increase of CRLBs (no significant change of CRLBs) was observed in spectra acquired with water suppression. For the clinical data, there were no substantial changes in the CRLB with water suppression. Because (1) the water suppression does not overall improve the quantification of the lipid resonances and (2) the MR spectrum without water suppression is always acquired for fat fraction calculation, the optimal data-acquisition strategy for liver MRS is to acquire only the MR spectrum without water suppression. For quantification of hepatic lipid resonances, it is advantageous to perform MR spectroscopy without water suppression in a clinical and preclinical scenario (at moderate fields) [14].

7.5. Management of Information in Neuroimaging

Participants: Michael Kain, Olivier Commowick, Elise Bannier, Inès Fakhfakh, Justine Guillaumont, Florent Leray, Yao Yao, Christian Barillot.

The major topic that is addressed in this period concern the sharing of data and processing tools in neuroimaging (through the "Programme d'Investissement d'Avenir" project such as OFSEP and FLI-IAM) which led to build a suitable architecture to share images and processing tools, started from the NeuroBase project (supported by the French Ministry of Research). Our overall goal within these projects is to set up a computer infrastructure to facilitate the sharing of neuroimaging data, as well as image processing tools, in a distributed and heterogeneous environment. These consortium gathered expertise coming from several complementary domains of expertise: image processing in neuroimaging, workflows and GRID computing, ontology development and ontology-based mediation. This enables a large variety of users to diffuse, exchange or reach neuroimaging information with appropriate access means, in order to be able to retrieve information almost as easily as if the data were stored locally by means of the "cloud computing" Storage as a Service (SaaS) concept. As an example, the Shanoir environment has been successfully deployed to the Neurinfo platform where it is routinely used to manage images of the research studies. It is also currently being deployed for two large projects: OFSEP ("Observatoire Français de la Sclérose en Plaques") where up to 30000 patients will be acquired on a ten years frame, and the Image Analysis and Management (IAM) node of the France Life Imaging national infrastructure (FLI-IAM). Our team fulfills multiple roles in this nation-wide FLI project. Christian Barillot

is the chair of the IAM node, Olivier Commowick is participating in the working group workflow and image processing and Michael Kain is the technical manager of the node. Apart from the team members, software solutions like medInria and Shanoir are part of the final software platform.

ALPINES Project-Team

7. New Results

7.1. Communication avoiding algorithms for dense linear algebra

Our group continues to work on algorithms for dense linear algebra operations that minimize communication. During this year we focused on improving the performance of communication avoiding QR factorization as well as designing algorithms for computing rank revealing and low rank approximations of dense and sparse matrices.

In [2] we discuss the communication avoiding QR factorization of a dense matrix. The standard algorithm for computing the QR decomposition of a tall and skinny matrix (one with many more rows than columns) is often bottlenecked by communication costs. The algorithm which is implemented in LAPACK, ScaLAPACK, and Elemental is known as Householder QR. For tall and skinny matrices, the algorithm works column-by-column, computing a Householder vector and applying the corresponding transformation for each column in the matrix. When the matrix is distributed across a parallel machine, this requires one parallel reduction per column. The TSQR algorithm, on the other hand, performs only one reduction during the entire computation. Therefore, TSQR requires asymptotically less inter-processor synchronization than Householder QR on parallel machines (TSQR also achieves asymptotically higher cache reuse on sequential machines). However, TSQR produces a different representation of the orthogonal factor and therefore requires more software development to support the new representation. Further, implicitly applying the orthogonal factor to the trailing matrix in the context of factoring a square matrix is more complicated and costly than with the Householder representation.

We show how to perform TSQR and then reconstruct the Householder vector representation with the same asymptotic communication efficiency and little extra computational cost. We demonstrate the high performance and numerical stability of this algorithm both theoretically and empirically. The new Householder reconstruction algorithm allows us to design more efficient parallel QR algorithms, with significantly lower latency cost compared to Householder QR and lower bandwidth and latency costs compared with Communication-Avoiding QR (CAQR) algorithm. Experiments on supercomputers demonstrate the benefits of the communication cost improvements: in particular, our experiments show substantial improvements over tuned library implementations for tall-and-skinny matrices. We also provide algorithmic improvements to the Householder QR and CAQR algorithms, and we investigate several alternatives to the Householder reconstruction algorithm that sacrifice guarantees on numerical stability in some cases in order to obtain higher performance.

In [4] we introduce CARRQR, a communication avoiding rank revealing QR factorization with tournament pivoting. Revealing the rank of a matrix is an operation that appears in many important problems as least squares problems, low rank approximations, regularization, nonsymmetric eigenproblems. In practice the QR factorization with column pivoting often works well, and it is widely used even if it is known to fail, for example on the so-called Kahan matrix. However in terms of communication, the QR factorization with column pivoting is sub-optimal with respect to lower bounds on communication. If the algorithm is performed in parallel, then typically the matrix is distributed over P processors by using a two-dimensional block cyclic partitioning. This is indeed the approach used in the `psgeqpf` routine from ScaLAPACK. At each step of the decomposition, the QR factorization with column pivoting finds the column of maximum norm and permutes it to the leading position, and this requires exchanging $O(n)$ messages, where n is the number of columns of the input matrix. For square matrices, when the memory per processor used is on the order of $O(n^2/P)$, the lower bound on the number of messages to be exchanged is $\Omega(\sqrt{P})$. The number of messages exchanged during the QR factorization with column pivoting is larger by at least a factor of n/\sqrt{P} than the lower bound.

In this paper we introduce CARRQR, a communication optimal (modulo polylogarithmic factors) rank revealing QR factorization based on tournament pivoting. The factorization is based on an algorithm that computes the decomposition by blocks of b columns (panels). For each panel, tournament pivoting proceeds in two steps. The first step aims at identifying a set of b candidate pivot columns that are as well-conditioned as possible. These columns are permuted to the leading positions, and they are used as pivots for the next b steps of the QR factorization. To identify the set of b candidate pivot columns, a tournament is performed based on a reduction operation, where at each node of the reduction tree b candidate columns are selected by using the strong rank revealing QR factorization. The idea of tournament pivoting has been first used to reduce communication in Gaussian elimination, an algorithm referred to as CALU.

We show that CARRQR reveals the numerical rank of a matrix in an analogous way to QR factorization with column pivoting (QRCP). Although the upper bound of a quantity involved in the characterization of a rank revealing factorization is worse for CARRQR than for QRCP, our numerical experiments on a set of challenging matrices show that this upper bound is very pessimistic, and CARRQR is an effective tool in revealing the rank in practical problems.

Our main motivation for introducing CARRQR is that it minimizes data transfer, modulo polylogarithmic factors, on both sequential and parallel machines, while previous factorizations as QRCP are communication sub-optimal and require asymptotically more communication than CARRQR. Hence CARRQR is expected to have a better performance on current and future computers, where communication is a major bottleneck that highly impacts the performance of an algorithm.

7.2. Algebraic preconditioners

Our work focused on the design of robust algebraic preconditioners and domain decomposition methods to accelerate the convergence of iterative methods.

In [5] we present a communication avoiding ILU0 preconditioner for solving large linear systems of equations by using iterative Krylov subspace methods. Recent research has focused on communication avoiding Krylov subspace methods based on so called s -step methods. However there is no communication avoiding preconditioner yet, and this represents a serious limitation of these methods. Our preconditioner allows to perform s iterations of the iterative method with no communication, through ghosting some of the input data and performing redundant computation. It thus reduces data movement by a factor of $3s$ between different levels of the memory hierarchy in a serial computation and between different processors in a parallel computation. To avoid communication, an alternating reordering algorithm is introduced for structured and unstructured matrices, that requires the input matrix to be ordered by using a graph partitioning technique such as kway or nested dissection. We show that the reordering does not affect the convergence rate of the ILU0 preconditioned system as compared to kway or nested dissection ordering, while it reduces data movement and should improve the expected time needed for convergence. In addition to communication avoiding Krylov subspace methods, our preconditioner can be used with classical methods such as GMRES or s -step methods to reduce communication.

7.3. A robust coarse space for Optimized Schwarz methods SORAS-GenEO-2

Optimized Schwarz methods (OSM) are very popular methods which were introduced by P.L. Lions for elliptic problems and Després for propagative wave phenomena. In [18], we have built a coarse space for which the convergence rate of the two-level method is guaranteed regardless of the regularity of the coefficients. We do this by introducing a symmetrized variant of the ORAS (Optimized Restricted Additive Schwarz) algorithm and by identifying the problematic modes using two different generalized eigenvalue problems instead of only one as for the ASM (Additive Schwarz method), BDD (balancing domain decomposition) or FETI (finite element tearing and interconnection) methods.

7.4. Time-dependent wave splitting and source separation

Starting from classical absorbing boundary conditions, we propose, in [17], a method for the separation of time-dependent scattered wave fields due to multiple sources or obstacles. In contrast to previous techniques, our method is local in space and time, deterministic, and also avoids a priori assumptions on the frequency spectrum of the signal. Numerical examples in two space dimensions illustrate the usefulness of wave splitting for time-dependent scattering problems.

7.5. Boundary integral formulations of wave scattering

We have continued to develop and further analyze new boundary integral formulation for wave scattering by complex objects.

In [13] we considered acoustic scattering of time-harmonic waves at objects composed of several homogeneous parts. Some of those may be impenetrable, giving rise to Dirichlet boundary conditions on their surfaces. We started from the second-kind boundary integral approach of [X. Claeys, and R. Hiptmair, and E. Spindler. A second-kind Galerkin boundary element method for scattering at composite objects. BIT Numerical Mathematics, 55(1):33-57, 2015] and extended it to this new setting. Based on so-called global multi-potentials, we derived variational second-kind boundary integral equations posed in $L^2(\Sigma)$, where Σ denotes the union of material interfaces. To suppress spurious resonances, we introduced a combined-field version (CFIE) of our new method. We conducted thorough numerical tests that highlighted the low and mesh-independent condition numbers of Galerkin matrices obtained with discontinuous piecewise polynomial boundary element spaces. They also confirmed competitive accuracy of the numerical solution in comparison with the widely used first-kind single-trace approach.

We spent much effort investigating the potentialities of multi-trace formulations in terms of domain decomposition. We considered multi-trace formulations in this perspective. Indeed Multi-Trace Formulations are based on a decomposition of the problem domain into subdomains, and thus domain decomposition solvers are of interest. The fully rigorous mathematical MTF can however be daunting for the non-specialist. In [12], we introduced MTFs on simple model problems using concepts familiar to researchers in domain decomposition. This allowed us to get a new understanding of MTFs and a natural block Jacobi iteration, for which we determined optimal relaxation parameters. We then showed how iterative multitrace formulation solvers are related to a well known domain decomposition method called optimal Schwarz method: a method which used Dirichlet to Neumann maps in the transmission condition. We finally showed that the insight gained from the simple model problem leads to remarkable identities for Calderón projectors and related operators, and the convergence results and optimal choice of the relaxation parameter we obtained is independent of the geometry, the space dimension of the problem, and the precise form of the spatial elliptic operator, like for optimal Schwarz methods. We confirmed this analysis with numerical experiments.

This work was extended in [10]. Considering pure transmission scattering problems in piecewise constant media, we derived an exact analytic formula for the spectrum of the corresponding local multi-trace boundary integral operators in the case where the geometrical configuration does not involve any junction point and all wave numbers equal. We deduced from this the essential spectrum in the case where wave numbers vary. Numerical evidences of these theoretical results were obtained in 2D.

Finally, in connection with boundary integral formulations, we extended the past work of [X. Claeys and R. Hiptmair, *Integral equations on multi-screens*. Integral Equations and Operator Theory, 77(2):167–197, 2013] where we had developed a framework for the analysis of boundary integral equations for acoustic scattering at so-called multi-screens, which are arbitrary arrangements of thin panels made of impenetrable material. In [3] we extended these considerations to boundary integral equations for electromagnetic scattering.

Viewing tangential multi-traces of vector fields from the perspective of quotient spaces we introduced the notion of single-traces and spaces of jumps. We also derived representation formulas and established key properties of the involved potentials and related boundary operators. Their coercivity were proved using a splitting of jump fields. Another new aspect emerged in the form of surface differential operators linking various trace spaces.

7.6. Asymptotic models for time harmonic wave propagation

Asymptotic models oriented toward more efficient numerical simulation methods have been investigated in three different directions.

In [8] we considered the Poisson equation in a domain with a small hole of size δ , and presented a simple numerical method, based on an asymptotic analysis, which allows to approximate robustly the far field of the solution as δ goes to zero without meshing the small hole. We proved the stability of the scheme and provide error estimates. This was confirmed with numerous numerical experiments illustrating the efficiency of the technique.

In [11] we considered a Laplace problem with Dirichlet boundary condition in a three dimensional domain containing an inclusion taking the form of a thin tube with small thickness. We proved convergence in operator norm of the resolvent of this problem as the thickness goes to 0, establishing that the perturbation on the resolvent induced by the inclusion is not greater than some (negative) power of the logarithm of the thickness. From this we deduced convergence of the eigenvalues of the perturbed operator toward the limit operator.

In [9] we investigated the eigenvalue problem $-\operatorname{div}(\sigma \nabla u) = \lambda u$ (\mathcal{P}) in a 2D domain Ω divided into two regions Ω_{\pm} . We were interested in situations where σ takes positive values on Ω_{+} and negative ones on Ω_{-} . Such problems appear in time harmonic electromagnetics in the modeling of plasmonic technologies. In a recent work [L. Chesnel, X. Claeys, and S.A. Nazarov. *A curious instability phenomenon for a rounded corner in presence of a negative material*. *Asymp. Anal.*, 88(1):43–74, 2014], we had highlighted an unusual instability phenomenon for the source term problem associated with (\mathcal{P}): for certain configurations, when the interface between the subdomains Ω_{\pm} presents a rounded corner, the solution may depend critically on the value of the rounding parameter. In [9] we explained this property studying the eigenvalue problem (\mathcal{P}). We provided an asymptotic expansion of the eigenvalues and prove error estimates. We established an oscillatory behaviour of the eigenvalues as the rounding parameter of the corner tends to zero. This work was ended with numerical illustrations.

7.7. New results related to FreeFem++

In [6], we consider a model of soil water and nutrient transport with plant root uptake. The geometry of the plant root system is explicitly taken into account in the soil model. We first describe our modeling approach. Then, we introduce an adaptive mesh refinement procedure enabling us to accurately capture the geometry of the root system and small-scale phenomena in the rhizosphere. Finally, we present a domain decomposition technique for solving the problems arising from the soil model as well as some numerical results.

In [15], we study an interface transport scheme of a two-phase flow of an incompressible viscous immiscible fluid. The problem is discretized by the characteristics method in time and finite elements in space. The interface is captured by the Level-Set function. Appropriate boundary conditions for the problem of mould filling are investigated, a new natural boundary condition under pressure effect for the transport equation is proposed and an algorithm for computing the solution is presented. Finally, numerical experiments show and validate the effectiveness of the proposed scheme.

ASAP Project-Team

6. New Results

6.1. Models and Theory of Distributed Systems

6.1.1. *Asynchronous Byzantine Systems: From Multivalued to Binary Consensus with $t < n/3$, $O(n^2)$ Messages, $O(1)$ Time, and no Signature*

Participant: Michel Raynal.

This work [39] presents a new algorithm that reduces multivalued consensus to binary consensus in an asynchronous message-passing system made up of n processes where up to t may commit Byzantine failures. This algorithm has the following noteworthy properties: it assumes $t < n/3$ (and is consequently optimal from a resilience point of view), uses $O(n^2)$ messages, has a constant time complexity, and does not use signatures. The design of this reduction algorithm relies on two new all-to-all communication abstractions. The first one allows the non-faulty processes to reduce the number of proposed values to c , where c is a small constant. The second communication abstraction allows each non-faulty process to compute a set of (proposed) values such that, if the set of a non-faulty process contains a single value, then this value belongs to the set of any non-faulty process. Both communication abstractions have an $O(n^2)$ message complexity and a constant time complexity. The reduction of multivalued Byzantine consensus to binary Byzantine consensus is then a simple sequential use of these communication abstractions. To the best of our knowledge, this is the first asynchronous message-passing algorithm that reduces multivalued consensus to binary consensus with $O(n^2)$ messages and constant time complexity (measured with the longest causal chain of messages) in the presence of up to $t < n/3$ Byzantine processes, and without using cryptography techniques. Moreover, this reduction algorithm tolerates message reordering by Byzantine processes.

This work was done in collaboration with Achour Mostefaoui from the LINA laboratory in Nantes.

6.1.2. *Atomic Read/Write Memory in Signature-free Byzantine Asynchronous Message-passing Systems*

Participant: Michel Raynal.

In this work [54] we designed a signature-free distributed algorithm which builds an atomic read/write shared memory on top of an n -process asynchronous message-passing system in which up to $t < n/3$ processes may commit Byzantine failures. From a conceptual point of view, this algorithm is designed to be as close as possible to the algorithm proposed by Attiya, Bar-Noy and Dolev (JACM 1995), which builds an atomic register in an n -process asynchronous message-passing system where up to $t < n/2$ processes may crash. The proposed algorithm is particularly simple. It does not use cryptography to cope with Byzantine processes, and is optimal from a t -resilience point of view ($t < n/3$). A read operation requires $O(n)$ messages, and a write operation requires $O(n^2)$ messages.

This work was done in collaboration with Achour Mostefaoui, Matoula Petrolia, and Claude Jard from the LINA laboratory in Nantes.

6.1.3. *Intrusion-Tolerant Broadcast and Agreement Abstractions in the Presence of Byzantine Processes*

Participant: Michel Raynal.

A process commits a Byzantine failure when its behavior does not comply with the algorithm it is assumed to execute. Considering asynchronous message-passing systems, this work [18] presents distributed abstractions, and associated algorithms, that allow non-faulty processes to correctly cooperate, despite the uncertainty created by the net effect of asynchrony and Byzantine failures. These abstractions are broadcast abstractions (namely, no-duplicity broadcast, reliable broadcast, and validated broadcast), and agreement abstraction (namely, consensus). While no-duplicity broadcast and reliable broadcast are well-known one-to-all communication abstractions, validated broadcast is a new all-to-all communication abstraction designed to address agreement problems. After having introduced these abstractions, we also presented an algorithm implementing validated broadcast on top of reliable broadcast. Then we presented two consensus algorithms, which are reductions of multivalued consensus to binary consensus. The first one is a generic algorithm, that can be instantiated with unreliable broadcast or no-duplicity broadcast, while the second is a consensus algorithm based on validated broadcast. Finally, a third algorithm is presented that solves the binary consensus problem. This algorithm is a randomized algorithm based on validated broadcast and a common coin. The presentation of all the abstractions and their algorithms is done incrementally. This work was done in collaboration with Achour Mostefaoui from the LINA laboratory in Nantes.

6.1.4. Minimal Synchrony for Asynchronous Byzantine Consensus

Participants: Michel Raynal, Zohir Bouzid.

Solving the consensus problem requires in one way or another that the underlying system satisfies some synchrony assumption. Considering an asynchronous message-passing system of n processes where (a) up to $t < n/3$ may commit Byzantine failures, and (b) each pair of processes is connected by two uni-directional channels (with possibly different timing properties), this work [24] investigates the synchrony assumption required to solve consensus, and presents a signature-free consensus algorithm whose synchrony requirement is the existence of a process that is an eventual $t+1$ bsource. Such a process p is a correct process that eventually has (a) timely input channels from t correct processes and (b) timely output channels to t correct processes (these input and output channels can connect p to different subsets of processes). As this synchrony condition was shown to be necessary and sufficient in the stronger asynchronous system model (a) enriched with message authentication, and (b) where the channels are bidirectional and have the same timing properties in both directions, it follows that it is also necessary and sufficient in the weaker system model considered in this work. In addition to the fact that it closes a long-lasting problem related to Byzantine agreement, a noteworthy feature of the proposed algorithm lies in its design simplicity, which is a first-class property.

This work was done in collaboration with Achour Mostefaoui from the LINA laboratory in Nantes.

6.1.5. Signature-Free Asynchronous Binary Byzantine Consensus with $t < n/3$, $O(n^2)$ Messages, and $O(1)$ Expected Time

Participant: Michel Raynal.

This work [17] is on broadcast and agreement in asynchronous message-passing systems made up of n processes, and where up to t processes may have a Byzantine Behavior. Its first contribution is a powerful, yet simple, all-to-all broadcast communication abstraction suited to binary values. This abstraction, which copes with up to $t < n/3$ Byzantine processes, allows each process to broadcast a binary value, and obtain a set of values such that (1) no value broadcast only by Byzantine processes can belong to the set of a correct process, and (2) if the set obtained by a correct process contains a single value v , then the set obtained by any correct process contains v . The second contribution of this work is a new round-based asynchronous consensus algorithm that copes with up to $t < n/3$ Byzantine processes. This algorithm is based on the previous binary broadcast abstraction and a weak common coin. In addition of being signature-free and optimal with respect to the value of t , this consensus algorithm has several noteworthy properties: the expected number of rounds to decide is constant; each round is composed of a constant number of communication steps and involves $O(n^2)$ messages; each message is composed of a round number plus a constant number of bits. Moreover, the algorithm tolerates message reordering by the adversary (i.e., the Byzantine processes). This work was done in collaboration with Achour Mostefaoui from the LINA laboratory in Nantes, and Hamouma Moumen from Université de Béjaïa.

6.1.6. Specifying Concurrent Problems: Beyond Linearizability and up to Tasks

Participants: Michel Raynal, Zohir Bouzid.

Tasks and objects are two predominant ways of specifying distributed problems. A task specifies for each set of processes (which may run concurrently) the valid outputs of the processes. An object specifies the outputs the object may produce when it is accessed sequentially. Each one requires its own implementation notion, to tell when an execution satisfies the specification. For objects linearizability is commonly used, while for tasks implementation notions are less explored. Sequential specifications are very convenient, especially important is the locality property of linearizability, which states that linearizable objects compose for free into a linearizable object. However, most well-known tasks have no sequential specification. Also, tasks have no clear locality property. This work [26] introduces the notion of interval-sequential object. The corresponding implementation notion of interval-linearizability generalizes linearizability. Interval-linearizability allows to specify any task. However, there are sequential one-shot objects that cannot be expressed as tasks, under the simplest interpretation of a task. We also showed that a natural extension of the notion of a task is expressive enough to specify any interval-sequential object.

This work was done in collaboration with Armando Castaneda and Sergio Rajsbaum from UNAM, Mexico.

6.1.7. Test-and-Set in Optimal Space

Participant: George Giakkoupis.

The test-and-set object is a fundamental synchronization primitive for shared memory systems. In [35] we address the number of registers (supporting atomic reads and writes) required to implement a one-shot test-and-set object in the standard asynchronous shared memory model with n processes. The best lower bound is $\log n - 1$ for obstruction-free and deadlock-free implementations, and recently a deterministic obstruction-free implementation using $O(\sqrt{n})$ registers was presented.

In [35] we close the gap between these existing upper and lower bounds by presenting a deterministic obstruction-free implementation of a one-shot test-and-set object from $\Theta(\log n)$ registers of size $\Theta(\log n)$ bits. Combining our obstruction-free algorithm with techniques from previous research, we also obtain a randomized wait-free test-and-set algorithm from $\Theta(\log n)$ registers, with expected step-complexity $\Theta(\log^* n)$ against the oblivious adversary. The core tool in our algorithm is the implementation of a deterministic obstruction-free *sifter* object, using only 6 registers. If k processes access a sifter, then when they have terminated, at least one and at most $\lfloor (2k + 1)/3 \rfloor$ processes return “win” and all others return “lose”.

This is a joint work with Maryam Helmi (U. of Calgary), Lisa Higham (U. of Calgary), and Philipp Woelfel (U. of Calgary), supported by the RADCON Inria Associate Team.

6.2. Graph and Probabilistic Algorithms

6.2.1. On the Quadratic Shortest Path Problem

Participant: Davide Frey.

Finding the shortest path in a directed graph is one of the most important combinatorial optimization problems, having applications in a wide range of fields. In its basic version, however, the problem fails to represent situations in which the value of the objective function is determined not only by the choice of each single arc, but also by the combined presence of pairs of arcs in the solution. In this work [40] we model these situations as a Quadratic Shortest Path Problem, which calls for the minimization of a quadratic objective function subject to shortest-path constraints. We prove strong NP-hardness of the problem and analyze polynomially solvable special cases, obtained by restricting the distance of arc pairs in the graph that appear jointly in a quadratic monomial of the objective function. Based on this special case and problem structure, we devise fast lower bounding procedures for the general problem and show computationally that they clearly outperform other approaches proposed in the literature in terms of its strength.

6.2.2. Tight Bounds on Vertex Connectivity Under Vertex Sampling

Participant: George Giakkoupis.

A fundamental result by Karger (SODA 1994) states that for any λ -edge-connected graph with n nodes, independently sampling each edge with probability $p = \Omega(\log n/\lambda)$ results in a graph that has edge connectivity $\Omega(\lambda p)$, with high probability. In [27] we prove the analogous result for vertex connectivity, when sampling vertices. We show that for any k -vertex-connected graph G with n nodes, if each node is independently sampled with probability $p = \Omega(\sqrt{\log n/k})$, then the subgraph induced by the sampled nodes has vertex connectivity $\Omega(kp^2)$, with high probability. This bound improves upon the recent results of Censor-Hillel et al. (SODA 2014) and is existentially optimal.

This is a joint work with Keren Censor-Hillel (Technion), Mohsen Ghaffari (MIT), Bernhard Haeupler (Carnegie Mellon U.), and Fabian Kuhn (U. of Freiburg).

6.3. Scalable Systems

6.3.1. *Being prepared in a sparse world: the case of KNN graph construction*

Participants: Anne-Marie Kermarrec, Nupur Mittal, Francois Taïani.

This work presents KIFF [41], a generic, fast and scalable KNN graph construction algorithm. KIFF directly exploits the bipartite nature of most datasets to which KNN algorithms are applied. This simple but powerful strategy drastically limits the computational cost required to rapidly converge to an accurate KNN solution, especially for sparse datasets. Our evaluation on a representative range of datasets show that KIFF provides, on average, a speed-up factor of 14 against recent state-of-the-art solutions while improving the quality of the KNN approximation by 18

This work was done in collaboration with Antoine Boutet from CNRS, Laboratoire Hubert Curien, Saint-Etienne, France.

6.3.2. *Cheap and Cheerful: Trading Speed and Quality for Scalable Social Recommenders*

Participants: Anne-Marie Kermarrec, François Taïani, Juan M. Tirado Martin.

Recommending appropriate content and users is a critical feature of on-line social networks. Computing accurate recommendations on very large datasets can however be particularly costly in terms of resources, even on modern parallel and distributed infrastructures. As a result, modern recommenders must generally trade-off quality and computational cost to reach a practical solution. This trade-off has however so far been largely left unexplored by the research community, making it difficult for practitioners to reach informed design decisions. In this work [37], we investigate to which extent the additional computing costs of advanced recommendation techniques based on supervised classifiers can be balanced by the gains they bring in terms of quality. In particular, we compare these recommenders against their unsupervised counterparts, which offer lightweight and highly scalable alternatives. We propose a thorough evaluation comparing 11 classifiers against 7 lightweight recommenders on a real Twitter dataset. Additionally, we explore data grouping as a method to reduce computational costs in a distributed setting while improving recommendation quality. We demonstrate how classifiers trained using data grouping can reduce their computing time by 6 while improving recommendations up to 22% when compared with lightweight solutions.

6.3.3. *Fast Nearest Neighbor Search*

Participants: Fabien André, Anne-Marie Kermarrec.

Nearest Neighbor (NN) search in high dimension is an important feature in many applications, such as multimedia databases, information retrieval or machine learning. Product Quantization (PQ) is a widely used solution which offers high performance, i.e., low response time while preserving a high accuracy. PQ represents high-dimensional vectors by compact codes. Large databases can therefore be stored in memory, allowing NN queries without resorting to slow I/O operations. PQ computes distances to neighbors using cache-resident lookup tables, thus its performance remains limited by (i) the many cache accesses that the algorithm requires, and (ii) its inability to leverage SIMD instructions available on modern CPUs.

To address these limitations, we designed a novel algorithm, PQ Fast Scan [19], that transforms the cache-resident lookup tables into small tables, sized to fit SIMD registers. This transformation allows (i) in-register lookups in place of cache accesses and (ii) an efficient SIMD implementation. PQ Fast Scan has the exact same accuracy as PQ, while having 4 to 6 times lower response time (e.g., for 25 million vectors, scan time is reduced from 74ms to 13ms).

This work was done in collaboration with Nicolas Le Scouarnec.

6.3.4. Holons: towards a systematic approach to composing systems of systems

Participants: Yérom-David Bromberg, François Taïani.

The world's computing infrastructure is increasingly differentiating into self-contained distributed systems with various purposes and capabilities (e.g. IoT installations, clouds, VANETs, WSNs, CDNs, . . .). Furthermore, such systems are increasingly being composed to generate systems of systems that offer value-added functionality. Today, however, system of systems composition is typically ad-hoc and fragile. It requires developers to possess an intimate knowledge of system internals and low-level interactions between their components. In this work [21], we outline a vision and set up a research agenda towards the generalised programmatic construction of distributed systems as compositions of other distributed systems. Our vision, in which we refer uniformly to systems and to compositions of systems as holons, employs code generation techniques and uses common abstractions, operations and mechanisms at all system levels to support uniform system of systems composition. We believe our holon approach could facilitate a step change in the convenience and correctness with which systems of systems can be built, and open unprecedented opportunities for the emergence of new and previously-unenvisaged distributed system deployments, analogous perhaps to the impact the mashup culture has had on the way we now build web applications.

This work was done in collaboration with Gordon Blair Geoff Coulson, and Yehia Elkhatib from Lancaster University (UK), Laurent Réveillère from University of Bordeaux / Labri, and Heverson Borba Ribeiro and Etienne Rivière from University of Neuchâtel (Switzerland).

6.3.5. Hybrid datacenter scheduling

Participant: Anne-Marie Kermarrec.

We address the problem of efficient scheduling of large clusters under high load and heterogeneous workloads. A heterogeneous workload typically consists of many short jobs and a small number of large jobs that consume the bulk of the cluster's resources.

Recent work advocates distributed scheduling to overcome the limitations of centralized schedulers for large clusters with many competing jobs. Such distributed schedulers are inherently scalable, but may make poor scheduling decisions because of limited visibility into the overall resource usage in the cluster. In particular, we demonstrate that under high load, short jobs can fare poorly with such a distributed scheduler.

We propose instead a new hybrid centralized/ distributed scheduler, called Hawk. In Hawk, long jobs are scheduled using a centralized scheduler, while short ones are scheduled in a fully distributed way. Moreover, a small portion of the cluster is reserved for the use of short jobs. In order to compensate for the occasional poor decisions made by the distributed scheduler, we propose a novel and efficient randomized work-stealing algorithm.

We evaluate Hawk using a trace-driven simulation and a prototype implementation in Spark. In particular, using a Google trace, we show that under high load, compared to the purely distributed Sparrow scheduler, Hawk improves the 50th and 90th percentile runtimes by 80% and 90% for short jobs and by 35% and 10% for long jobs, respectively. Measurements of a prototype implementation using Spark on a 100-node cluster confirm the results of the simulation. This work has been done in the context of the Inria/epfl research center and in collaboration with Pamela delgado, Florin Dinu and Willy Zwaenepoel from EPFL and published in Usenix ATC in 2015 [30].

6.3.6. Out-of-core KNN Computation

Participants: Nitin Chiluka, Anne-Marie Kermarrec, Javier Olivares.

This work proposes a novel multi threading approach to compute KNN on large datasets by leveraging both disk and main memory efficiently. The main rationale of our approach is to minimize random accesses to disk, maximize sequential access to data and efficient usage of only a fraction of the available memory. This approach is evaluated by comparing its performance with a fully in-memory implementation of KNN, in terms of execution time and memory consumption. This multithreading approach outperforms the in-memory baseline in all cases when the large dataset does not fit in memory.

6.3.7. *Scaling Out Link Prediction with SNAPLE*

Participants: Anne-Marie Kermarrec, François Taïani, Juan M. Tirado Martin.

A growing number of organizations are seeking to analyze extra large graphs in a timely and resource-efficient manner. With some graphs containing well over a billion elements, these organizations are turning to distributed graph-computing platforms that can scale out easily in existing data-centers and clouds. Unfortunately such platforms usually impose programming models that can be ill suited to typical graph computations, fundamentally undermining their potential benefits. In this work [38], we consider how the emblematic problem of link-prediction can be implemented efficiently in gather-apply-scatter (GAS) platforms, a popular distributed graph-computation model. Our proposal, called Snaple, exploits a novel highly-localized vertex scoring technique, and minimizes the cost of data flow while maintaining prediction quality. When used within GraphLab, Snaple can scale to very large graphs that a standard implementation of link prediction on GraphLab cannot handle. More precisely, we show that Snaple can process a graph containing 1.4 billions edges on a 256 cores cluster in less than three minutes, with no penalty in the quality of predictions. This result corresponds to an over-linear speedup of 30 against a 20-core standalone machine running a non-distributed state-of-the-art solution.

6.3.8. *Similitude: Decentralised Adaptation in Large-Scale P2P Recommenders*

Participants: Davide Frey, Anne-Marie Kermarrec, Pierre-Louis Roman, François Taïani.

Decentralised recommenders have been proposed to deliver privacy-preserving, personalised and highly scalable on-line recommendations. Current implementations tend, however, to rely on a hard-wired similarity metric that cannot adapt. This constitutes a strong limitation in the face of evolving needs. In this work [33], we propose a framework to develop dynamically adaptive decentralized recommendation systems. Our proposal supports a decentralised form of adaptation, in which individual nodes can independently select, and update their own recommendation algorithm, while still collectively contributing to the overall system's mission.

This work was done in collaboration with Christopher Maddock and Andreas Mauthe (Univ. of Lancaster, UK).

6.3.9. *Transactional Memory Recommenders*

Participant: Anne-Marie Kermarrec.

The Transactional Memory (TM) paradigm promises to greatly simplify the development of concurrent applications. This led, over the years, to the creation of a plethora of TM implementations delivering wide ranges of performance across workloads. Yet, no universal TM implementation fits each and every workload. In fact, the best TM in a given workload can reveal to be disastrous for another one. This forces developers to face the complex task of tuning TM implementations, which significantly hampers the wide adoption of TMs. In this work, we address the challenge of automatically identifying the best TM implementation for a given workload. Our proposed system, ProteusTM, hides behind the TM interface a large library of implementations. Under the hood, it leverages an innovative, multi-dimensional online optimization scheme, combining two popular machine learning techniques: Collaborative Filtering and Bayesian Optimization. We integrated ProteusTM in GCC and demonstrated its ability to switch TM implementations and adapt several configuration parameters (e.g., number of threads). We extensively evaluated ProteusTM, obtaining average performance 3% less than the optimal, and gains up to 100 over static alternatives.

This work has been done in collaboration with Rachid Guerraoui from EPFL, Diego Didona Nuno Diegues, Ricardo Neves and Paolo Romano from INESC, Lisboa) and will be published in ASPLOS 2016 [31].

6.3.10. *Want to scale in centralized systems? Think P2P*

Participants: Anne-Marie Kermarrec, François Taïani.

Peer-to-peer (P2P) systems have been widely researched over the past decade, leading to highly scalable implementations for a wide range of distributed services and applications. A P2P system assigns symmetric roles to machines, which can act both as client and server. This distribution of responsibility alleviates the need for any central component to maintain a global knowledge of the system. Instead, each peer takes individual decisions based on a local and limited knowledge of the rest of the system, providing scalability by design. While P2P systems have been successfully applied to a wide range of distributed applications (multicast, routing, caches, storage, pub-sub, video streaming), with some highly visible successes (Skype, Bitcoin), they tend to have fallen out of fashion in favor of a much more cloud-centric vision of the current Internet. We think this is paradoxical, as cloud-based systems are themselves large-scale, highly distributed infrastructures. They reside within massive, densely interconnected datacenters, and must execute efficiently on an increasing number of machines, while dealing with growing volumes of data. Today even more than a decade ago, large-scale systems require scalable designs to deliver efficient services. In this work [16] we argue that the local nature of P2P systems is key for scalability regardless whether a system is eventually deployed on a single multi-core machine, distributed within a data center, or fully decentralized across multiple autonomous hosts. Our claim is backed by the observation that some of the most scalable services in use today have been heavily influenced by abstractions and rationales introduced in the context of P2P systems. Looking to the future, we argue that future large-scale systems could greatly benefit from fully decentralized strategies inspired from P2P systems. We illustrate the P2P legacy through several examples related to Cloud Computing and Big Data, and provide general guidelines to design large-scale systems according to a P2P philosophy.

6.3.11. *WebGC: Browser-based gossiping*

Participants: Raziel Carvajal Gomez, Davide Frey, Anne-Marie Kermarrec.

The advent of browser-to-browser communication technologies like WebRTC has renewed interest in the peer-to-peer communication model. However, the available WebRTC code base still lacks important components at the basis of several peer-to-peer solutions. Through a collaboration with Mathieu Simonin from the Inria SED in the context of the Brow2Brow ADT project, we started to tackle this problem by proposing WebGC, a library for gossip-based communication between web browsers. Due to their inherent scalability, gossip-based, or epidemic protocols constitute a key component of a large number of decentralized applications. WebGC thus represents an important step towards their wider spread. We demonstrated the final version of the library at WISE 2015 [53].

6.4. Privacy in User Centric Applications

6.4.1. *Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!*

Participants: Davide Frey, Anne-Marie Kermarrec, Antoine Rault.

Whether we are shopping for an interesting book or selecting a movie to watch, the chances are that a recommendation system will help us decide what we want. Recommendation systems collect information about our own preferences, compare them to those of other users, and provide us with suggestions on a variety of topics. But is the information gathered by a recommendation system safe from potential attackers, be them other users, or companies that access the recommendation system? And above all, can service providers protect this information while still providing effective recommendations? In this work, we analyze the effect of Sybil attacks on collaborative-filtering recommendation systems, and discuss the impact of different similarity metrics in the trade-off between recommendation quality and privacy. Our results, on a state-of-the-art recommendation framework and on real datasets show that existing similarity metrics exhibit a wide range of behaviors in the presence of Sybil attacks. Yet, they are all subject to the same trade off: Sybil resilience for recommendation quality. We therefore propose and evaluate a novel similarity metric that combines the best of both worlds: a low RMSE score with a prediction accuracy for Sybil users of only a few percent. A preliminary version of this work was published at EuroSec 2015 [32].

This work was done in collaboration with Antoine Boutet, and Rachid Guerraoui.

6.4.2. Decentralized view prediction for global content placement

Participants: Stéphane Delbruel, Davide Frey, François Taïani.

A large portion of today's Internet traffic originates from streaming and video services. Storing, indexing, and serving these videos is a daily engineering challenge that requires increasing amounts of efforts and infrastructures. One promising direction to improve video services consists in predicting at upload time where and when a new video might be viewed, thereby optimizing placement and caching decisions. Implementing such a prediction service in a scalable manner poses significant technical challenges. In this work [28], we address these challenges in the context of a decentralized storage system consisting of set-top boxes or end nodes. Specifically, we propose a novel data placement algorithm that exploits information about the tags associated with existing content, such as videos, and uses it to infer the number of views that newly uploaded content will have in each country.

6.4.3. Distance-Based Differential Privacy in Recommenders

Participant: Anne-Marie Kermarrec.

The upsurge in the number of web users over the last two decades has resulted in a significant growth of online information. This information growth calls for recommenders that personalize the information proposed to each individual user. Nevertheless, personalization also opens major privacy concerns. We designed D2P, a novel protocol that ensures a strong form of differential privacy, which we call distance-based differential privacy, and which is particularly well suited to recommenders. D2P avoids revealing exact user profiles by creating altered profiles where each item is replaced with another one at some distance. We evaluate D2P analytically and experimentally on MovieLens and Jester datasets and compare it with other private and non-private recommenders. This work has been done in the context of the Web-Alter-ego Google focused award and in collaboration with Rachid guerraoui, Rhicheck Patra and Masha Taziki from EPFL and published in PVLVB in 2015 [15].

6.4.4. Privacy-Conscious Information Diffusion in Social Networks

Participants: George Giakkoupis, Arnaud Jégou, Anne-Marie Kermarrec, Nupur Mittal.

This work presents a distributed algorithm – Riposte [47], for information dissemination in social networks which guarantees to preserve the privacy of its users. RIPOSTE ensures that information spreads widely if and only if a large fraction of users find it interesting, and this is done in a “privacy-conscious” manner, namely without revealing the opinion of any individual user. Whenever an information item is received by a user, RIPOSTE decides to either forward the item to all the user's neighbors, or not to forward it to anyone. The decision is randomized and is based on the user's (private) opinion on the item, as well as on an upper bound s on the number of user's neighbors that have not received the item yet.

This work was done in collaboration with Rachid Guerraoui from EPFL, Switzerland.

6.4.5. Hide & Share: Landmark-based Similarity for Private knn Computation

Participants: Davide Frey, Anne-Marie Kermarrec, Antoine Rault, François Taïani.

Computing k-nearest-neighbor graphs constitutes a fundamental operation in a variety of data-mining applications. As a prominent example, user-based collaborative-filtering provides recommendations by identifying the items appreciated by the closest neighbors of a target user. As this kind of applications evolve, they will require KNN algorithms to operate on more and more sensitive data. This has prompted researchers to propose decentralized peer-to-peer KNN solutions that avoid concentrating all information in the hands of one central organization. Unfortunately, such decentralized solutions remain vulnerable to malicious peers that attempt to collect and exploit information on participating users.

In this work [22], we seek to overcome this limitation by proposing H&S (Hide & Share), a novel landmark-based similarity mechanism for decentralized KNN computation. Landmarks allow users (and the associated peers) to estimate how close they lay to one another without disclosing their individual profiles.

We evaluate H&S in the context of a user-based collaborative-filtering recommender with publicly available traces from existing recommendation systems. We show that although landmark-based similarity does disturb similarity values (to ensure privacy), the quality of the recommendations is not as significantly hampered. We also show that the mere fact of disturbing similarity values turns out to be an asset because it prevents a malicious user from performing a profile reconstruction attack against other users, thus reinforcing users' privacy. Finally, we provide a formal privacy guarantee by computing the expected amount of information revealed by H&S about a user's profile.

This work was done in collaboration with Antoine Boutet from the University of St. Etienne, and with Jingjing Wang and Rachid Guerraoui from EPFL, Switzerland.

ASCOLA Project-Team

6. New Results

6.1. Highlights of the year

Nicolas Tabareau has been awarded a starting grant from the European Research Council (ERC), the most prestigious type of research projects of the European Union for young researchers. From 2015–2020 he will pursue research on “CoqHoTT: Coq for Homotopy Type Theory.”

In the domain of resource management notably for Cloud infrastructures, the team has produced several very visible results. These include contributions to popular and new simulation tools and platforms [17], [27], [28] as well as new techniques for the energy-efficient execution of Cloud applications [15].

On the topics of software composition and programming languages, the team has, among others, two remarkable results: a new notion of effect capabilities and corresponding monadic analysis techniques [14] as well as the first comprehensive survey of domain-specific aspect languages [13].

6.2. Programming Languages

Participants: Walid Bengerbit, Ronan-Alexandre Cherrueau, Rémi Douence, Hervé Grall, Florent Marchand de Kerchove de Denterghem, Jacques Noyé, Jean-Claude Royer, Mario Südholt.

6.2.1. Formal Methods, logics and type theory

This year we have proposed “Gradual Certified Programming” as a bridge between type-based expressive proofs and programming languages, have extended previous type theories by new homotopy-based means, and have introduced “effect capabilities” to control monad-based effects in Haskell.

6.2.1.1. Gradual Certified Programming in Coq

Expressive static typing disciplines are a powerful way to achieve high-quality software. However, the adoption cost of such techniques should not be under-estimated. Just like gradual typing allows for a smooth transition from dynamically-typed to statically-typed programs, it seems desirable to support a gradual path to certified programming. We have explored gradual certified programming in Coq [33], providing the possibility to postpone the proofs of selected properties, and to check “at runtime” whether the properties actually hold. Casts can be integrated with the implicit coercion mechanism of Coq to support implicit cast insertion à la gradual typing. Additionally, when extracting Coq functions to mainstream languages, our encoding of casts supports lifting assumed properties into runtime checks. Much to our surprise, it is not necessary to extend Coq in any way to support gradual certified programming. A simple mix of type classes and axioms makes it possible to bring gradual certified programming to Coq in a straightforward manner.

6.2.1.2. Homotopy Hypothesis in Type Theory

In classical homotopy theory, the homotopy hypothesis asserts that the fundamental omega-groupoid construction induces an equivalence between topological spaces and weak omega-groupoids. In the light of Voevodsky’s univalent foundations program, which puts forward an interpretation of types as topological spaces, we have considered the question of transposing the homotopy hypothesis to type theory [16]. Indeed such a transposition could stand as a new approach to specifying higher inductive types. Since the formalization of general weak omega-groupoids in type theory is a difficult task, we have only taken a first step towards this goal, which consists in exploring a shortcut through strict omega-categories. The first outcome is a satisfactory type-theoretic notion of strict omega-category, which has hsets of cells in all dimensions. For this notion, defining the ‘fundamental strict omega-category’ of a type seems out of reach. The second outcome is an ‘incoherently strict’ notion of type-theoretic omega-category, which admits arbitrary types of cells in all dimensions. These are the ‘wild’ omega-categories of the title. They allow the definition of a ‘fundamental wild omega-category’ map, which leads to our (partial) homotopy hypothesis for type theory (stating an adjunction, not an equivalence). All of our results have been formalized in the Coq proof assistant. Our formalization makes systematic use of the machinery of coinductive types.

6.2.1.3. Effect Capabilities For Haskell

Computational effects complicate the tasks of reasoning about and maintaining software, due to the many kinds of interferences that can occur. While different proposals have been formulated to alleviate the fragility and burden of dealing with specific effects, such as state or exceptions, there is no prevalent robust mechanism that addresses the general interference issue. Building upon the idea of capability-based security, we have proposed effect capabilities [14] as an effective and flexible manner to control monadic effects and their interferences. Capabilities can be selectively shared between modules to establish secure effect-centric coordination. We have further refined capabilities with type-based permission lattices to allow fine-grained decomposition of authority. An implementation of effect capabilities in Haskell has been done, using type classes to establish a way to statically share capabilities between modules, as well as to check proper access permissions to effects at compile time.

6.2.1.4. Correct Refactoring Tools

Most integrated development environments provide refactoring tools. However, these tools are often unreliable. As a consequence, developers have to test their code after applying an automatic refactoring.

Refactoring tools for industrial languages are difficult to test and verify. We have developed a refactoring operation for C programs (renaming of global variables) for which we have proved that it preserves the set of possible behaviors of the transformed programs [39]. That proof of correctness relies on the operational semantics provided by CompCert C in Coq. We have also proved some properties of the transformation which are used to establish properties of a composed refactoring operations.

6.2.2. Language Mechanisms

This year we have contributed new results on domain-specific aspect languages, concurrent event-based programming, model transformations as well as the relationship between functional and constraint programming. Furthermore, we have proposed language support for the definition and enforcement of security properties, in particular related to the accountability of service-based systems, see Sec. 6.3 .

6.2.2.1. Domain-Specific Aspect Languages

Domain-Specific Aspect Languages (DSALs) are Domain-Specific Languages (DSLs) designed to express crosscutting concerns. Compared to DSLs, their aspectual nature greatly amplifies the language design space. In the context of the Associate Team RAPIDS/REAL, we have structured this space in order to shed light on and compare the different domain-specific approaches to deal with crosscutting concerns [13]. We have reported on a corpus of 36 DSALs covering the space, discussed a set of design considerations and provided a taxonomy of DSAL implementation approaches. This work serves as a frame of reference to DSAL and DSL researchers, enabling further advances in the field, and to developers as a guide for DSAL implementations.

6.2.2.2. Concurrent Event-Based Programming

The advanced concurrency abstractions provided by the Join calculus overcome the drawbacks of low-level concurrent programming techniques. However, with current approaches, the coordination logic involved in complex coordination schemas is still fragmented. In [11], Jurgen Van Ham presents JEScala, a language that captures coordination schemas in a more expressive and modular way by leveraging a seamless integration of an advanced event system with join abstractions. The implementation of joins-based state machines is discussed with alternative faster implementations made possible through a domain specific language. Event monitors are introduced as a way of synchronizing event handling and building concurrent event-based applications from sequential event-based parts.

6.2.2.3. Model Lazy Transformation

The Object Constraint Language (OCL) is a central component in modeling and transformation languages such as the Unified Modeling Language (UML), the Meta Object Facility (MOF), and Query View Transformation (QVT). OCL is standardized as a strict functional language. We have proposed a lazy evaluation strategy for OCL [36]. This lazy evaluation semantics is beneficial in some model-driven engineering scenarios for speeding up the evaluation times for very large models, simplifying expressions on models by using infinite

data structures (e.g., infinite models) and increasing the reusability of OCL libraries. We have implemented the approach on the ATL virtual machine EMFTVM. This is a joint work with the Inria team Atlanmod.

6.2.2.4. *Composition Mechanisms for Constraints Generalization*

Structural time series (pattern for sequences of values) can be described with numerous automata-based constraints. In [12], we describe a large family of constraints for structural time series by means of function composition. We formalize the patterns using finite transducers. Based on that description, we automatically synthesize automata with accumulators, as well as constraint checkers. The description scheme not only unifies the structure of the existing 30 time-series constraints, but also leads to over 600 new constraints, with more than 100,000 lines of synthesized code. This is a joint work with the Inria team Tasc.

6.3. Software Composition

Participants: Walid Benghrabit, Ronan-Alexandre Cherrueau, Rémi Douence, Hervé Grall, Jean-Claude Royer, Mario Südholt.

6.3.1. *Constructive Security*

Nowadays we are witnessing the wide-spread use of cloud services. As a result, more and more end-users (individuals and businesses) are using these services for achieving their electronic transactions (shopping, administrative procedures, B2B transactions, etc.). In such scenarios, personal data is generally flowing between several entities and end-users need (i) to be aware of the management, processing, storage and retention of personal data, and (ii) to have necessary means to hold service providers accountable for the usage of their data. Usual preventive security mechanisms are not adequate in a world where personal data can be exchanged on-line between different parties and/or stored at multiple jurisdictions. Accountability becomes a necessary principle for the trustworthiness of open computer systems. It regards the responsibility and liability for the data handling performed by a computer system on behalf of an organization. In case of misconduct (e.g. security breaches, personal data leak, etc.), accountability should imply remediation and redress actions, as in the real life.

In 2015, we have contributed two main results: first, techniques for the logic-based definition, analysis and verification of accountability properties; second, a new framework for the compositional definition of privacy-properties and their type-based enforcement.

6.3.1.1. *Logic-based accountability properties*

We have proposed a framework for the representation of accountability policies [37]. This framework comes with two novel accountability policy languages; the Abstract Accountability Language (AAL), which is devoted to the representation of preferences/obligations in an human readable fashion, and a concrete one for the mapping to concrete enforceable policies. Our efforts have focused on a formal foundation for the AAL language and some applications.

We have also introduced an approach to assist the design of accountable applications [21]. In particular, we consider an application's abstract component design and we introduce a logical approach allowing various static verification. This approach offers effective means to early check the design and the behavior of an application and its offered/required services. We motivate our work with a realistic use case coming from the A4Cloud project and validate our proposal with experiments using the TSPASS theorem prover. This prover is competitive with other model-checkers and sat solvers and we gain a more abstract approach than with our previous experiment with a model-checker. It makes also easier the link with end users, for instance privacy officers.

To give a formal foundation of the AAL language we define a translation into first-order temporal logic [20]. We introduce a formula to interpret accountability and a natural criterion to achieve the accountability compliance of two clauses. We continue to apply it to an health care system taking into account data privacy features, data transfers and location processing. We demonstrate few heuristics to speed up the resolution time and to assist in conflict detection. Tool support (AccLab) has been provided to support editing, checking and proving AAL clauses.

6.3.1.2. Composition of Privacy-Enforcement Techniques

Today's large-scale computations, e.g., in the Cloud, are subject to a multitude of risks concerning the divulging and ownership of private data. Privacy risks are mainly addressed using a large variety of encryption-based techniques. We have proposed a compositional approach for the declarative and correct composition of privacy-preserving applications in the Cloud [22], [38]. Our approach provides language support for the compositional definition of encryption-based and fragmentation-based privacy-preserving algorithms. This language comes equipped with a set of laws that allows us to verify privacy properties. We have provided implementation support in Scala that ensures certain privacy properties by construction using advanced features of Scala's type system.

6.3.2. Modular systems

6.3.2.1. Modularity for Javascript Interpreters.

With an initial motivation based on the security of web applications written in JavaScript, we have provided new techniques for the instrumentation of an interpreter for a dynamic analysis as a crosscutting concern [31]. We have defined the instrumentation problem — an extension to the expression problem with a focus on modifying interpreters. We have then shown how we can instrument an interpreter for a simple language using only the bare language features provided by JavaScript.

6.4. Cloud applications and infrastructures

Participants: Frederico Alvares, Simon Dupont, Md Sabbir Hasan, Adrien Lebre, Thomas Ledoux, Jonathan Lejeune, Guillaume Le Louët, Jean-Marc Menaud, Jonathan Pastor, Mario Südholt.

In 2015, we have provided solutions for Cloud-based and distributed programming, virtual environments and data centers.

6.4.1. Cloud and distributed programming

6.4.1.1. Cloud elasticity

Cloud Computing has provided important new means for the capacity management of resources. The elasticity and the economy of scale are the intrinsic elements that differentiate it from traditional computing paradigm.

A good capacity planning method is a necessary factor but not sufficient to fully exploit Cloud elasticity. In [26], we propose innovative policies for resource management to achieve the optimal balance between capacity and quality of Cloud services. The main idea is to finely control the scalability and the termination of virtual machines with respect to several criteria such as the lifecycle of the instances (e.g. initialization time) or their cost. The approach was evaluated on an Amazon EC2 cluster. Experimental results illustrate the soundness of the proposed approach and the impact of scalability/termination resource policies: a cost saving of as much as 30% can be achieved with a minimal number of violations, as small as 1%.

In order to improve Cloud elasticity, we advocate that the software layer can take part in the elasticity process as the overhead of software reconfiguration can be usually considered negligible compared to infrastructural costs. Thanks to this extra level of elasticity, we are able to define cloud reconfigurations that enact elasticity in both the software and infrastructure layers. In [23], we present an autonomic approach to manage cloud elasticity in a cross-layered manner. First, we enhance cloud elasticity with the software elasticity model. Then, we describe how our autonomic cloud elasticity model relies on the dynamic selection of elasticity tactics. We present an experimental analysis of a subset of those elasticity tactics under different scenarios in order to provide insights on strategies that could drive the autonomic selection of the proper tactics to be applied.

6.4.1.2. Service-level agreement for the Cloud

Quality-of-service and SLA guarantees are among the major challenges of cloud-based services. In [18], we first present a new cloud model called SLAaaS — SLA aware Service. SLAaaS considers QoS levels and SLA as first class citizens of cloud-based services. This model is orthogonal to other SaaS, PaaS, and IaaS cloud models, and may apply to any of them. More specifically, we make three contributions: (i) we provide a domain-specific language that allows to define SLA constraints in cloud services; (ii) we present a general control-theoretic approach for managing cloud service SLA; (iii) we apply our approach to MapReduce, locking, and e-commerce services.

6.4.1.3. Distributed multi-resource allocation

Generalized distributed mutual exclusion algorithms allow processes to concurrently access a set of shared resources. However, they must ensure an exclusive access to each resource. In order to avoid deadlocks, many of them are based on the strong assumption of a prior knowledge about conflicts between processes' requests. Some other approaches, which do not require such a knowledge, exploit broadcast mechanisms or a global lock, degrading message complexity and synchronization cost. We propose in [29] [41] a new solution for shared resources allocation which reduces the communication between non-conflicting processes without a prior knowledge of processes conflicts. Performance evaluation results show that our solution improves resource use rate by a factor up to 20 compared to a global lock based algorithm.

6.4.2. Virtualization and data centers

In 2015, we have produced results and tools for the simulation of large-scale distributed algorithms, notably VM scheduling algorithms, have contributed new abstractions for storage systems and have devised new means for the introspection of Cloud infrastructures.

6.4.2.1. SimGrid / VMPlaceS

We have developed VMPlaceS [28], a framework providing programming support for the definition of VM placement algorithms, execution support for their simulation at large scales, as well as new means for their trace-based analysis. VMPlaceS enables, in particular, the investigation of placement algorithms in the context of numerous and diverse real-world scenarios. To illustrate relevance of such a tool, we evaluated three different classes of virtualization environments: centralized, hierarchical and fully distributed placement algorithms. We showed that VMPlaceS facilitates the implementation and evaluation of variants of placement algorithms. The corresponding experiments have provided the first systematic results comparing these algorithms in environments including up to one thousand of nodes and ten thousands of VMs in most cases.

While such a number is already valuable and although we finalized the virtualization abstractions in SimGrid [17], we are in touch with the core developers in order to improve the code of VMPlaceS with the ultimate objective of addressing infrastructures up to 100K physical machines and 1 Millions virtual machines over a period of one day.

The current version of VMPlaceS is available on a public git repository :<http://beyondtheclouds.github.io/VMPlaceS/>.

6.4.2.2. Storage abstractions within the SimGrid framework

With the recent data deluge, storage is becoming the most important resource to master in modern computing infrastructures. Dimensioning and assessing the performance of storage systems are challenges for which simulation constitutes a sound approach. Unfortunately, only a few existing simulators of large scale distributed computing systems go beyond providing merely a notion of storage capacity. In 2015, we contributed to the SimGrid efforts toward the simulation of such systems [27]. Concretely, we characterized the performance behavior of several types of disks to derive a first model of storage resource. This model has been integrated within the SimGrid framework available under the LGPL license (<http://simgrid.gforge.inria.fr>).

6.4.2.3. Cloud Introspection

Cloud Computing has become a new technical and economic model for many IT companies. By virtualizing services, it allows for a more flexible management of datacenters capacities. However, its elasticity and its flexibility led to the explosion of virtual environments to manage. It's common for a system administrator to manage several hundreds or thousands virtual machines. Without appropriate tools, this administration task may be impossible to achieve.

We propose in [32] a decision support tool to detect virtual machines with atypical behavior. Virtual machines whose behavior is different from other VMs running in the data center are tagged as atypicals. Our analysis tool is based on a specific partitioning algorithm which identifies VM behaviors. This tool has been validated in production environments and is used by several companies.

To collect finer metrics (for security, energy management etc.), VM introspection an agent can be installed in a VM to intrusively supervise it or the hypervisor can be used to non-intrusively recover the introspection metrics. In the case of intrusive introspection, the agent installed on the VM operating system will retrieve a set of information related to the operating system operation. However, the installation of an agent in the virtual machine increases the cost of deploying the virtual machine and its resource consumption. The Virtual Machine Introspection (VMI) at the hypervisor level (non intrusively) offer a complete, consistent and untainted view of the VM state. This solution allows an isolation of the VMI mechanism from the guest OS, while allowing monitoring and modifying any state of the VM.

We have also provided a comprehensive summary on VM introspection techniques [25]. Existing VMI techniques are analyzed with respect to their approach to closing the "semantic gap" between the (low level) information provided by the hypervisor and the input to the security analysis.

Finally, we have introduced an extension to LibVmi to detect and monitor a process resource consumption inside a VM from the hypervisor [34]. This extension monitor process cpu and ram resources without probe. This extension can detect abusive cpu resource usage and atypical ram utilization. This fine monitoring system can be used in many context (security, power consumption, fault tolerance).

6.5. Green IT

Participants: Simon Dupont, Md Sabbir Hasan, Thomas Ledoux, Jonathan Lejeune, Guillaume Le Louët, Jean-Marc Menaud.

In 2015, we have provided new models and solutions for the energy-optimal execution of cloud applications in data centers.

6.5.1. Renewable energy

With the emergence of the Future Internet and the emergence of new IT models such as cloud computing, the usage of data centers (DC) and consequently their power consumption increase dramatically. Besides the ecological impact, the energy consumption is a predominant criteria for DC providers since it determines the daily cost of their infrastructure. As a consequence, power management becomes one of the main challenges for DC infrastructures and more generally for large-scale distributed systems.

6.5.1.1. Renewable energy for data centers

We have presented the EPOC project which focuses on optimizing the energy consumption of mono-site DCs connected to the regular electrical grid and to renewable energy sources [19]. A first challenge in this context consists in developing a (for users) transparent distributed system that enables energy-proportional computations from the system to service-oriented levels. The second challenge addresses the corresponding energy issues through collaborative measurements and energy-optimizing actions inside infrastructure-software stack, more precisely between applications and resource management systems. This approach must manage Service Level Agreement (SLA) constraints by striving for the best trade-off between energy cost (from the regular electric grid), its availability (from renewable energy sources), and service degradation (from application reconfiguration issues to job suspension ones). The third challenge embarks pursues energy efficient optical networks as key enablers of the future internet and cloud-networking service deployment through the convergence of optical infrastructure with the upper network layers.

The second challenge is more precisely describe in [30]. In this paper we present PIKA, a framework aiming at reducing the Brownian energy consumption (ie. from non renewable energy sources), and improving the usage of renewable energy for mono-site data centers. PIKA exploits jobs with slack periods, and executes and suspends them depending on the available renewable energy supply. By consolidating the virtual machines (VMs) on the physical servers, PIKA adjusts the number of powered-on servers in order for the overall energy consumption to match the renewable energy supply. Using simulations driven by real-world workloads and solar power traces, we demonstrate that PIKA consumes 41% less Brownian energy and increases 35.3% renewable energy integration ratio in comparison with the baseline algorithm from the literature.

6.5.1.2. Energy monitoring

We have designed SensorScript, a Business-Oriented Domain-Specific Language for Sensor Networks [24], [35]. In smart grids, or more generally the Internet of Things, many research work has been performed on the whole chain, from communication sensors to big data management, through communication middlewares. Few of this work have addressed the problem of gathered data access. In fact, being able, as a system administrator, to manipulate and gather data collected from a set of sensors in a simple and efficient way represents an essential need.

To address this issue, the solution we considered consists of a multi-context modeling for raw data, in the form of a multi-tree: a directed acyclic graph consisting of multiple intricate trees, each of them describing a hierarchy corresponding to a given use context. The objectives are to provide not only a means to rationalize users needs before writing queries, but also to offer a domain-specific language (DSL) which takes advantage of the multi-tree modeling to simplify the experience of pre-identified users that query data.

6.5.1.3. Green SLA and virtualization of green energy

The demand for energy-efficient services is increasing considerably as people are getting more environmentally-conscious in order to build a sustainable society. The main challenge for Cloud providers is to manage Green SLA (Service Level Agreement) constraints for their customers while satisfying their business objectives, such as maximizing profits by lowering expenditure for so-called green (renewable) energy. Since, Green SLA needs to be proposed based on the presence of green energy, the intermittent nature of renewable sources makes it difficult to be achieved. In response, we propose a scheme for green energy management based on three contributions [15]: i) we introduce the concept of virtualization of green energy to address the uncertainty of green energy availability, ii) we extend the Cloud Service Level Agreement (CSLA) language to support Green SLA by introducing two new threshold parameters and iii) we introduce algorithms for Green SLA which leverage the concept of virtualization of green energy to provide interval-specific Green SLA. We have conducted experiments with real workload profiles from PlanetLab and server power model from SPECpower to demonstrate that Green SLA can be successfully established and satisfied without incurring higher cost.

ATLANMODELS Team

7. New Results

7.1. Reverse Engineering & Evolution

Model Driven Reverse Engineering (MDRE), with its applications on software modernization or tool evolution for example, is a discipline in which model-based principles and techniques are used to treat various kinds of (sometimes very large) existing systems. In the continuity the work started several years ago, AtlanMod has been working actively on this research area this year again. The main contributions are the following:

- In the context of the ARTIST FP7 project, the work has been continued on reusing (and extending accordingly) MoDisco and several of its components to provide the Reverse Engineering support required within the project (and more particularly in the context of the use cases provided by our industrial partners). This has been an important year for the team in this project since it successfully ended in November 2015 after final review at the European Commission. At conceptual-level, the proposed overall approach (as a main result of the ARTIST project) and the main lessons we have learned from its application to concrete industrial scenarios have been published and promoted to a large and high-level audience [11]. The ARTIST project in itself, the various research aspects it addresses and the offered technical solutions have also been presented to the Modeling community [22]. At tooling-level, several (MoDisco-based) model discovery components from Java and SQL have been enhanced while made available as part of a second version of the official ARTIST OS Release ⁰. A promising work has also been started on studying deeper the automated discovery of behavioral aspects of software applications, notably by working on a pragmatic mapping between a programming language (Java) and a modeling language (the OMG fUML standard) that focuses on executable aspects.
- To facilitate the understanding of existing software applications via the different models describing them, a significant work has been performed related to providing a generic support for dealing with viewpoints and views expressed on set of possibly heterogeneous and large models. To this intent, and directly capitalizing on the work performed in the TEAP FUI project that ended by the end of 2014, the EMF Views prototype has been significantly refined and enhanced with a ViewPoint Definition Language (the VPDL domain-specific language having a SQL-like syntax) notably [18]. Based on this same model viewpoints/views approach, and more particularly on its underlying (meta)model virtualization support, the general problem of lightweight (meta)model extension has been studied more deeply in the context of our work within the MoNoGe FUI project (national). This has already resulted in a corresponding prototype and a DSL for expressing metamodel extensions [17]. Within the coming year, the plan is to continue further this global work on model viewpoints/views in a software understanding and evolution context.
- Software development projects are notoriously complex and difficult to deal with. Several support tools have been introduced in the past decades to ease the development activities such as issue tracking, code review and Source Control Management (SCM) systems. While such tools efficiently track the evolution of a given aspect of the project (e.g., issues or code), they provide just a partial view of the software project and they often lack of querying mechanisms beyond basic support (e.g., command line, simple gui). This is particularly true for projects that rely on Git, one of the most popular SCM systems. Nowadays many tools are built on top of it, however, they do not complement Git with query functionalities and currently none of them proposes a mechanism that unifies the project information scattered in such different tools. In [28], we propose a conceptual schema for Git and an approach that, given a Git repository, exports its data to a relational database in order to (1) promote data integration with other existing Git-based tools relying on databases

⁰<http://www.artist-project.eu/tools-of-toolbox/193>

and (2) provide query functionalities expressed through standard SQL syntax. To ensure efficiency, our approach comes with an incremental propagation mechanism that refreshes the database content with the latest modifications.

7.2. MDE Scalability

The increasing number of companies embracing MDE methods and tools have exceeded the limits of the current model-based technologies, presenting scalability issues while facing the growing complexity of their data. Since further research and development is imperative in order to maintain MDE techniques as relevant as they are in less complex contexts, we have focused our research in three axes, (i) scalable persistence solutions, (ii) scalable model transformation engines, and (iii) testing of large scale distributed systems.

In [33], we introduce and evaluate a map-based persistence model for MDE tools. We use this model to build a transparent persistence layer for modeling tools, on top of a map-based database engine. The layer can be plugged into the Eclipse Modeling Framework, lowering execution times and memory consumption levels of other existing approaches. Empirical tests are performed based on a typical industrial scenario, model-driven reverse engineering, where very large software models originate from the analysis of massive code bases. The layer is freely distributed and can be immediately used for enhancing the scalability of any existing Eclipse Modeling tool. We learned that—in terms of performance—typical model-access APIs, with fine-grained methods that only allow for one-step-navigation queries, do not benefit from complex relational or graph-based data structures. Much better results are potentially obtained by optimized low-level data structures, like hash-tables, which guarantee low and constant access times. Additional features that may be of interest in scenarios where performance is not an issue (such as versioning and transactional support provided by CDO) have not been considered. In [32] we extend our persistent mechanism to distributed environments by presenting NeoEMF/HBase, a model-persistence backend for the Eclipse Modeling Framework (EMF) built on top of the Apache HBase data store. Model distribution is hidden from client applications, that are transparently provided with the model elements they navigate. Access to remote model elements is decentralized, avoiding the bottleneck of a single access point. The persistence model is based on key-value stores that allow for efficient on-demand model persistence.

Once we develop a high-performance and distributed persistence mechanism for very-large models, we can exploit it to run high-performance computing over such models. One of the central operations in MDE is rule-based model transformation (MT). It is used to specify manipulation operations over structured data coming in the form of model graphs. However, being based on computationally expensive operations like subgraph isomorphism, MT tools are facing issues on both memory occupancy and execution time while dealing with the increasing model size and complexity. One way to overcome these issues is to exploit the wide availability of distributed clusters in the Cloud for the distributed execution of MT. In [24] and [23], we propose an approach to automatically distribute the execution of model transformations written in a popular MT language, ATL, on top of a well-known distributed programming model, MapReduce. We show how the execution semantics of ATL can be aligned with the MapReduce computation model. We describe the extensions to the ATL transformation engine to enable distribution, and we experimentally demonstrate the scalability of this solution in a reverse-engineering scenario.

Another fundamental operation in MDE is model querying. The Object Constraint Language (OCL) is the standard query language proposed by OMG and is a central component in other modeling and transformation languages such as the Unified Modeling Language (UML), the Meta Object Facility (MOF), and Query View Transformation (QVT). OCL is standardized as a strict functional language. In [34], we propose a lazy evaluation strategy for OCL. We argue that a lazy evaluation semantics is beneficial in some model-driven engineering scenarios for: i) lowering evaluation times on very large models; ii) simplifying expressions on models by using infinite data structures (e.g., infinite models); iii) increasing the reusability of OCL libraries. We implement the approach on the ATL virtual machine EMFTVM.

Finally an important class of operations in MDE is bidirectional (i.e. reversible) computation. Especially bidirectional model transformation is a key technology when two models that can change over time have to be kept constantly consistent with each other. In Hidaka et al. we clarify and visualize the space of design

choices for bidirectional transformations from an MDE point of view, in the form of a feature model. The selected list of existing approaches are characterized by mapping them to the feature model. Then the feature model is used to highlight some unexplored research lines in bidirectional transformations, especially in the scalability of such systems.

7.3. Software Quality

We initiated a new line of research in order to investigate Novelty Search (NS) for the automatic generation of test data, in collaboration with the DiverSE team. Our goal is to explore the huge space of test data within the input domain. In this approach, we select test data based on a novelty score showing how different they are compared to all other solutions evaluated so far [25], [26].

AVALON Project-Team

7. New Results

7.1. Energy Efficiency of Large Scale Distributed Systems

Participants: Laurent Lefèvre, Daniel Balouek Thomert, Eddy Caron, Radu Carpa, Marcos Dias de Assunção, Jean-Patrick Gelas, Olivier Glück, Jean-Christophe Mignot, Violaine Villebonnet.

7.1.1. Energy efficient Core Networks

This work [8], [43] seeks to improve the energy efficiency of backbone networks by providing an intra-domain Software Defined Network (SDN) approach to selectively turn off a subset of links. To do this, we designed an energy-aware traffic engineering technique for reducing energy consumption in backbone networks. Energy-efficient traffic engineering was analysed in previous work, but none addressed implementation challenges of their solutions. We showed that ignoring to test the feasibility of techniques can lead to bad estimations and unstable solutions. We proposed the STREETE framework (Segment Routing based Energy Efficient Traffic Engineering) that represents an online method to switch some links off/on dynamically according to the network load. We have implemented a working prototype in the OMNET++ simulator. Networks are progressively using centralised architecture, and SDN is increasingly utilised in data centre networks. We believe that SDN may be extended to backbone networks. The implemented solution shows that SDN may also be a good means for reducing the energy consumption of network devices. Compared to previous work, in this work we used the SPRING protocol to improve the stability of energy-efficient traffic engineering solutions. To the best of our knowledge, this is the first work proposing the use of SPRING to improve the energy efficiency of backbone networks. The flexibility of this routing protocol is well suited to frequent route changes that happen when we switch links off and on. Moreover, this protocol can be easily applied to SDN solutions. Using simulations, we showed that as much as 44% of links can be switched off to save energy in real backbone networks. Even greedy techniques can easily approach the maximum reduction in the amount of energy consumed. In fact, the bottleneck in terms of energy efficiency in energy-aware traffic engineering is the connectivity constraint. We performed a stress test of our solution under rapidly increasing traffic patterns and showed that more work must be done in the domain of switching links back on: a field which has received little attention from the research community.

7.1.2. Energy proportionality in HPC systems

Energy savings are among the most important topics concerning Cloud and HPC infrastructures nowadays. Servers consume a large amount of energy, even when their computing power is not fully utilized. These static costs represent quite a concern, mostly because many datacenter managers are over-provisioning their infrastructures compared to the actual needs. This results in a high part of wasted power consumption. In this work [19], [47], we proposed the BML (“Big, Medium, Little”) infrastructure, composed of heterogeneous architectures, and a scheduling framework dealing with energy proportionality. We introduce heterogeneous power processors inside datacenters as a way to reduce energy consumption when processing variable workloads. Our framework brings an intelligent utilization of the infrastructure by dynamically executing applications on the architecture that suits their needs, while minimizing energy consumption. Our first validation process focuses on distributed stateless web servers scenario and we analyze the energy savings achieved through energy proportionality. This research activity is performed with the collaboration of Sepia Team (IRIT, Toulouse) through the co-advising of Violaine Villebonnet.

7.1.3. Energy-Aware Server Provisioning

Several approaches to reduce the power consumption of datacenters have been described in the literature, most of which aim to improve energy efficiency by trading off performance for reducing power consumption. However, these approaches do not always provide means for administrators and users to specify how they want to explore such trade-offs. This work [27] provides techniques for assigning jobs to distributed resources, exploring energy efficient resource provisioning. We use middleware-level mechanisms to adapt resource allocation according to energy-related events and user-defined rules. A proposed framework enables developers, users and system administrators to specify and explore energy efficiency and performance trade-offs without detailed knowledge of the underlying hardware platform. Evaluation of the proposed solution under three scheduling policies shows gains of 25% in energy-efficiency with minimal impact on the overall application performance. We also evaluate reactivity in the adaptive resource provisioning. This approach has been applied in the Nuage research project [26].

7.1.4. Virtual Home Gateway

About 80-90% of the energy in today's wireline networks is consumed in the access network, including about 10 to 30W per user being dissipated mostly by the customer premises equipment (CPE). Home gateway is a popular equipment deployed at the end of networks and supporting a set of heterogeneous services (data, phone, television, multimedia, security services). These gateways and associated services can be difficult to deploy and maintain for customers. These gateways are difficult to manage for network operators and consume a lot of energy. We explore the technical solutions to reduce the complexity and energy impact of such equipments by moving services to some external dedicated and shared facilities of network operators. This result is a joint work between Avalon team (J.P. Gelas, L. Lefevre) and Addis Abeba University (M. Tsibie and T. Assefa). This research has been demonstrated in the GreenTouch final celebration event in New York (June 2015).

7.2. MPI Application and Storage System Simulation

Participants: Frédéric Suter, Laurent Pouilloux.

7.2.1. Scalable Off-line Simulation of MPI Applications

Analyzing and understanding the performance behavior of parallel applications on parallel computing platforms is a long-standing concern in the High Performance Computing community. When the targeted platforms are not available, simulation is a reasonable approach to obtain objective performance indicators and explore various hypothetical scenarios. In the context of applications implemented with the Message Passing Interface, two simulation methods have been proposed, on-line simulation and off-line simulation, both with their own drawbacks and advantages.

We proposed in [9] an off-line simulation framework, i.e., one that simulates the execution of an application based on event traces obtained from an actual execution. The main novelty of this work, when compared to previously proposed off-line simulators, is that traces that drive the simulation can be acquired on large, distributed, heterogeneous, and non-dedicated platforms. As a result the scalability of trace acquisition is increased, which is achieved by enforcing that traces contain no time-related information. Moreover, our framework is based on an state-of-the-art scalable, fast, and validated simulation kernel.

Such off-line analysis faces scalability issues for acquiring, storing, or replaying large event traces. Then, in [10], we combined our framework with another, specialized in the production of compact traces, to capitalize on their respective strengths while alleviating several of their limitations. We showed that the combined framework affords levels of scalability that are beyond that achievable by either one of the two individual frameworks.

7.2.2. Simulation of Storage Elements

Storage is an essential component of distributed computing infrastructures, *i.e.*, clusters, grids, clouds, data centers, or supercomputers, to cope with the tremendous increase in scientific data production and the ever-growing need for data analysis and preservation. Understanding the performance of a storage subsystem or dimensioning it properly is an important concern for which simulation can help by allowing for fast, fully repeatable, and configurable experiments for arbitrary hypothetical scenarios. However, most simulation frameworks tailored for the study of distributed systems offer no or little abstractions or models of storage resources.

In [34], we detailed the extension of SimGrid with storage simulation capacities. We first defined the required abstractions and propose a new API to handle storage components and their contents in SimGrid-based simulators. Then we characterized the performance of the fundamental storage component that are disks and derive models of these resources. Finally we listed several concrete use cases of storage simulations in clusters, grids, clouds, and data centers for which the proposed extension would be beneficial.

7.3. MapReduce Computations on Hybrid Distributed Computations Infrastructures

Participants: Gilles Fedak, Julio Anjos, Anthony Simonet.

In this section we report on our efforts to provide MapReduce Computing environments on Hybrid infrastructures, *i.e.* composed of Desktop Grids and Cloud computing environments.

Cloud computing has increasingly been used as a platform for running large business and data processing applications. Although clouds have become extremely popular, when it comes to data processing, their use incurs high costs. Conversely, Desktop Grids, have been used in a wide range of projects, and are able to take advantage of the large number of resources provided by volunteers, free of charge. Merging cloud computing and desktop grids into a hybrid infrastructure can provide a feasible low-cost solution for big data analysis. Although frameworks like MapReduce have been devised to exploit commodity hardware, their use in a hybrid infrastructure raise some challenges due to their large resource heterogeneity and high churn rate.

7.3.1. BIGHybrid - A Toolkit for Simulating MapReduce in Hybrid Infrastructures

In [20], we introduced BIGHybrid, a toolkit that is used to simulate MapReduce in hybrid environments. Its main goal is to provide a framework for developers and system designers that can enable them to address the issues of Hybrid MapReduce. In this paper, we described the framework which simulates the assembly of two existing middleware: BitDew- MapReduce for Desktop Grids and Hadoop-BlobSeer for Cloud Computing. The experimental results that are included in this work demonstrate the feasibility of our approach.

7.3.2. HybridMR: a New Approach for Hybrid MapReduce Combining Desktop Grid and Cloud Infrastructures

In [18], we proposed a novel MapReduce computation model in hybrid computing environment called HybridMR. Using this model, high performance cluster nodes and heterogeneous desktop PCs in Internet or Intranet can be integrated to form a hybrid computing environment. In this way, the computation and storage capability of large-scale desktop PCs can be fully utilized to process large-scale datasets. HybridMR relies on a hybrid distributed file system called HybridDFS, and a time-out method has been used in HybridDFS to prevent volatility of desktop PCs, and file replication mechanism is used to realize reliable storage. A new node priority-based fair scheduling (NPBFS) algorithm has been developed in HybridMR to achieve both data storage balance and job assignment balance by assigning each node a priority through quantifying CPU speed, memory size and I/O bandwidth. Performance evaluation results showed that the proposed hybrid computation model not only achieves reliable MapReduce computation, reduces task response time and improves the performance of MapReduce, but also reduces the computation cost and achieves a greener computing mode.

7.3.3. *D³ -MapReduce: Towards MapReduce for Distributed and Dynamic Data Sets*

So far MapReduce has been mostly designed for batch processing of bulk data. The ambition of D³-MapReduce, presented in [32], is to extend the MapReduce programming model and propose efficient implementation of this model to: i) cope with distributed data sets, i.e. that span over multiple distributed infrastructures or stored on network of loosely connected devices; ii) cope with dynamic data sets, i.e. which dynamically change over time or can be either incomplete or partially available. In this paper, we draw the path towards this ambitious goal. Our approach leverages Data Life Cycle as a key concept to provide MapReduce for distributed and dynamic data sets on heterogeneous and distributed infrastructures. We first reported on our attempts at implementing the MapReduce programming model for Hybrid Distributed Computing Infrastructures (Hybrid DCIs). We present the architecture of the prototype based on BitDew, a middleware for large scale data management, and Active Data, a programming model for data life cycle management. Second, we outlined the challenges in term of methodology and present our approaches based on simulation and emulation on the Grid'5000 experimental testbed. We conducted performance evaluations and compare our prototype with Hadoop, the industry reference MapReduce implementation. We presented our work in progress on dynamic data sets that has lead us to implement an incremental MapReduce framework. Finally, we discussed our achievements and outline the challenges that remain to be addressed before obtaining a complete D³-MapReduce environment.

7.3.4. *Availability and Network-Aware MapReduce Task Scheduling over the Internet.*

MapReduce offers an ease-of-use programming paradigm for processing large datasets. In our previous work, we have designed a MapReduce framework called BitDew-MapReduce for desktop grid and volunteer computing environment, that allows nonexpert users to run data-intensive MapReduce jobs on top of volunteer resources over the Internet. However, network distance and resource availability have great impact on MapReduce applications running over the Internet. To address this, an availability and network-aware MapReduce framework over the Internet is proposed in [38]. Simulation results show that the MapReduce job response time could be decreased by 27.15%, thanks to Naive Bayes Classifier-based availability prediction and landmark-based network estimation.

7.4. Managing Big Data Life Cycle

Participants: Gilles Fedak, Anthony Simonet.

7.4.1. *Active Data - Enabling Smart Data Life Cycle Management for Large Distributed Scientific Data Sets*

The Big Data challenge consists in managing, storing, analyzing and visualizing these huge and ever growing data sets to extract sense and knowledge. As the volume of data grows exponentially, the management of these data becomes more complex in proportion. A key point is to handle the complexity of the data life cycle, i.e. the various operations performed on data: transfer, archiving, replication, deletion, etc. Indeed, data-intensive applications span over a large variety of devices and e-infrastructures which implies that many systems are involved in data management and processing. In [17], we proposed Active Data, a programming model to automate and improve the expressiveness of data management applications. We first define the concept of data life cycle and introduce a formal model that allows to expose data life cycle across heterogeneous systems and infrastructures. The Active Data programming model allows code execution at each stage of the data life cycle: routines provided by programmers are executed when a set of events (creation, replication, transfer, deletion) happen to any data. We implement and evaluate the model with four use cases: a storage cache to Amazon-S3, a cooperative sensor network, an incremental implementation of the MapReduce programming model and automated data provenance tracking across heterogeneous systems. Altogether, these scenarios illustrate the adequateness of the model to program applications that manage distributed and dynamic data sets. We also show that applications that do not leverage on data life cycle can still benefit from Active Data to improve their performances.

7.4.2. Using Active Data to Provide Smart Data Surveillance to E-Science Users

Modern scientific experiments often involve multiple storage and computing platforms, software tools, and analysis scripts. The resulting heterogeneous environments make data management operations challenging, the significant number of events and the absence of data integration makes it difficult to track data provenance, manage sophisticated analysis processes, and recover from unexpected situations. Current approaches often require costly human intervention and are inherently error prone. The difficulties inherent in managing and manipulating such large and highly distributed datasets also limits automated sharing and collaboration. In [37], we study a real world e-Science application involving terabytes of data, using three different analysis and storage platforms, and a number of applications and analysis processes. We demonstrate that using a specialized data life cycle and programming model, Active Data, we can easily implement global progress monitoring, and sharing, recover from unexpected events, and automate a range of tasks.

7.4.3. SMART: An Application Framework for Real Time Big Data Analysis on Heterogeneous Cloud Environments.

The amount of data that human activities generate poses a challenge to current computer systems. Big data processing techniques are evolving to address this challenge, with analysis increasingly being performed using cloud-based systems. Emerging services, however, require additional enhancements in order to ensure their applicability to highly dynamic and heterogeneous environments and facilitate their use by Small & Medium-sized Enterprises (SMEs). Observing this landscape in emerging computing system development, this work presents Small & Medium-sized Enterprise Data Analytic in Real Time (SMART) for addressing some of the issues in providing compute service solutions for SMEs. SMART offers a framework for efficient development of Big Data analysis services suitable to small and medium-sized organizations, considering very heterogeneous data sources, from wireless sensor networks to data warehouses, focusing on service composability for a number of domains. In [62], we presented the basis of this proposal and preliminary results on exploring application deployment on hybrid infrastructure.

7.5. Desktop Grid Computing

Participants: Gilles Fedak, Anthony Simonet.

7.5.1. Multi-Criteria and Satisfaction Oriented Scheduling for Hybrid Distributed Computing Infrastructures

Assembling and simultaneously using different types of distributed computing infrastructures (DCI) like Grids and Clouds is an increasingly common situation. Because infrastructures are characterized by different attributes such as price, performance, trust, greenness, the task scheduling problem becomes more complex and challenging. In [15], we presented the design for a fault-tolerant and trust-aware scheduler, which allows to execute Bag-of-Tasks applications on elastic and hybrid DCI, following user-defined scheduling strategies. Our approach, named Promethee scheduler, combines a pull-based scheduler with multi-criteria Promethee decision making algorithm. Because multi-criteria scheduling leads to the multiplication of the possible scheduling strategies, we proposed SOFT, a methodology that allows to find the optimal scheduling strategies given a set of application requirements. The validation of this method is performed with a simulator that fully implements the Promethee scheduler and recreates an hybrid DCI environment including Internet Desktop Grid, Cloud and Best Effort Grid based on real failure traces. A set of experiments shows that the Promethee scheduler is able to maximize user satisfaction expressed accordingly to three distinct criteria: price, expected completion time and trust, while maximizing the infrastructure useful employment from the resources owner point of view. Finally, we present an optimization which bounds the computation time of the Promethee algorithm, making realistic the possible integration of the scheduler to a wide range of resource management software.

7.5.2. Synergy of Volunteer Measurements and Volunteer Computing for Effective Data Collecting, Processing, Simulating and Analyzing on a Worldwide Scale

The paper [31] concerns the hype idea of Citizen Science and the related paradigm shift: to go from the passive “volunteer computing” to other volunteer actions like “volunteer measurements” under guidance of scientists. They can be carried out by ordinary people with standard computing gadgets (smartphone, tablet, etc.) and the various standard sensors in them. Here the special attention is paid to the system of volunteer scientific measurements to study air showers caused by cosmic rays. The technical implementation is based on integration of data about registered night flashes (by radiometric software) in shielded camera chip, synchronized time and GPS-data in ordinary gadgets: to identify night air showers of elementary particles; to analyze the frequency and to map the distribution of air showers in the densely populated cities. The project currently includes the students of the National Technical University of Ukraine KPI, which are compactly located in Kyiv city and contribute their volunteer measurements. The technology would be very effective for other applications also, especially if it will be automated (e.g., on the basis of XtremWeb or/and BOINC technologies for distributed computing) and used in some small area with many volunteers, e.g. in local communities (Corporate/Community Crowd Computing).

7.5.3. Towards an Environment for doing Data Science that runs in Browsers

In [25], we proposed a path for doing Data Science using browsers as computing and data nodes. This novel idea is motivated by the cross-fertilized fields of desktop grid computing, data management in grids and clouds, Web technologies such as Nosql tools, models of interactions and programming models in grids, cloud and Web technologies. We propose a methodology for the modeling, analyzing, implementation and simulation of a prototype able to run a MapReduce job in browsers. This work allows to better understand how to envision the big picture of Data Science in the context of the Javascript language for programming the middleware, the interactions between components and browsers as the operating system. We explain what types of applications may be impacted by this novel approach and, from a general point of view how a formal modeling of the interactions serves as a general guidelines for the implementation. Formal modeling in our methodology is a necessary condition but it is not sufficient. We also make round-trips between the modeling and the Javascript or used tools to enrich the interaction model that is the key point, or to put more details into the implementation. It is the first time to the best of our knowledge that Data Science is operating in the context of browsers that exchange codes and data for solving computational and data intensive programs. Computational and data intensive terms should be understood according to the context of applications that we think to be suitable for our system.

7.5.4. E-Fast & CloudPower: Towards High Performance Technical Analysis for Small Investors.

About 80% of the financial market investors fail, the main reason for this being their poor investment decisions. Without advanced financial analysis tools and the knowledge to interpret the analysis, the investors can easily make irrational investment decisions. Moreover, investors are challenged by the dynamism of the market and a relatively large number of indicators that must be computed. In this paper we propose E-Fast, an innovative approach for on-line technical analysis for helping small investors to obtain a greater efficiency on the market by increasing their knowledge. The E-Fast technical analysis platform prototype relies on High Performance Computing (HPC), allowing to rapidly develop and extensively validate the most sophisticated finance analysis algorithms. In [36], we aim at demonstrating that the E-Fast implementation, based on the CloudPower HPC infrastructure, is able to provide small investors a realistic, low-cost and secure service that would otherwise be available only to the large financial institutions. We describe the architecture of our system and provide design insights. We present the results obtained with a real service implementation based on the Exponential Moving Average computational method, using CloudPower and Grid5000 for the computations' acceleration. We also elaborate a set of interesting challenges emerging from this work, as next steps towards high performance technical analysis for small investors.

7.6. HPC Component Model

Participants: Hélène Coullon, Vincent Lanore, Christian Perez, Jérôme Richard.

7.6.1. 3D FFT and L^2C

We have completed the work started in 2014. To harness the computing power of supercomputers, HPC application algorithms have to be adapted to the underlying hardware. This is a costly and complex process which requires handling many algorithm variants. In [23], we studied the ability of the component model L^2C to express and handle the variability of HPC applications. The goal is to ease application adaptation. Analysis and experiments are done on a 3D-FFT use case. Results show that L^2C , and components in general, offer a generic and simple handling of 3D-FFT variants while obtaining performance close to well-known libraries

7.6.2. Multi-Stencil DSL in L^2C

As high performance architectures evolve continuously to be more powerful, such architectures also usually become more difficult to use efficiently. As a scientist is not a low level and high performance programming expert, Domain Specific Languages (DSLs) are a promising solution to automatically and efficiently write high performance codes. However, if DSLs ease programming for scientists, maintainability and portability issues are transferred from scientists to DSL designers. This work [44] has dealt with an approach to improve maintainability and programming productivity of DSLs through the generation of a component-based parallel runtime. To study it, we have designed a DSL for multi-stencil programs, that is evaluated on a real-case of shallow water equations implemented with L^2C .

7.6.3. Reconfigurable HPC component model

High-performance applications whose structure changes dynamically during execution are extremely complex to develop, maintain and adapt to new hardware. Such applications would greatly benefit from easy reuse and separation of concerns which are typical advantages of component models. Unfortunately, no existing component model is both HPC-ready (in terms of scalability and overhead) and able to easily handle dynamic reconfiguration. In [33], we aimed at addressing performance, scalability and programmability by separating locking and synchronization concerns from reconfiguration code. To this end, we propose directMOD, a component model which provides on one hand a flexible mechanism to lock subassemblies with a very small overhead and high scalability, and on the other hand a set of well-defined mechanisms to easily plug various independently-written reconfiguration components to lockable subassemblies. We evaluate both the model itself and a C++/MPI implementation called directL2C .

7.6.4. Towards a Task-Component Model

In [24], we propose a first model that aims at combining both component models and task based models such as StarPU. Component models bring many good software engineering properties such as code re-use while task based models seems to be very efficient to exploit recent hardware such as SMP, manycore, or GPGPUs. This work evaluates a proof-of-concepts only considering SMP nodes.

7.7. Security for Virtualization and Clouds

Participants: Eddy Caron, Arnaud Lefray.

7.7.1. Security and placement

We have proposed a solution for placement-based security and client-centric security. Even with perfect information flow control mechanisms, virtualized environments are still sensitive to silent information leakage, that is covert channels, due to shared hardware resources. We have proposed a fine-grained placement based on the client's security properties to tackle this issue. The client submits an application i.e., a graph of VMs, and information flow rules defining the acceptable risk. Due to the lack of usable covert channel metric to qualify an acceptable risk, we have proposed a new information leakage metric. As covert channels exploit microarchitecture flaws, we have integrated the specificity of NUMA allocation schemes in our placement algorithm.

7.7.2. Security and logic language

Besides, the main issue with existing security languages is the ability to formally guarantee the required property. On the one hand, security policies described in a natural language have quite ambiguous semantics. On the other hand, a formal language or logic provides clear syntax and semantics. Moreover, existing mechanisms are dedicated to secure specific type of entities (e.g., VM, Service, Data, VNet). Therefore, the problem is to have a formal definition of security properties and proven procedures to transform the end-user's global security properties into multiple local properties enforceable by several local mechanisms. For these reasons, we proposed a logic language called IF-PLTL (Information Flow Past Linear Time Logic). Our logic is dedicated to controlling the propagation of information i.e., direct and indirect information flows. As these information flows cannot be obtained directly, we have explained their construction from low-level observable events. Security decisions are naturally expressed according to past actions. Accordingly, IF-PLTL is based on the past fragment of LTL. In addition to using IF-PLTL to transform properties, we have proposed a dynamic monitor that can enforce the full expressivity of IF-PLTL even if its complexity (in time and space) would incur a high overhead in practice.

7.8. Autonomic Middleware Deployment using Self-Stabilization

Participants: Eddy Caron, Maurice Faye.

Dynamic nature of distributed architecture is a major challenge to avail the benefits of distributed computing. An effective solution to deal with this dynamic nature is to implement a self-adaptive mechanism to sustain the distributed architecture. Self-adaptive systems can autonomously modify their behavior at run-time in response to changes in their environment. This capability may be included in the software systems at design time or later by external mechanisms. We have created a self-adaptive algorithm for the DIET middleware. Once the middleware is deployed, it can detect a set of events which indicate an unstable deployment state. When an event is detected, some instructions are executed to handle the event. We have designed a simulator to have a deeper insights of our proposed self-adaptive algorithm.

CIDRE Project-Team

7. New Results

7.1. Intrusion detection

7.1.1. Alert Correlation in Distributed Systems

In large systems, multiple (host and network) Intrusion Detection Systems (IDS) and many sensors are usually deployed. They continuously and independently generate notifications (event's observations, warnings and alerts). To cope with this amount of collected data, alert correlation systems have to be designed. An alert correlation system aims at exploiting the known relationships between some elements that appear in the flow of low level notifications to generate high semantic meta-alerts. The main goal is to reduce the number of alerts returned to the security administrator and to allow a higher level analysis of the situation. However, producing correlation rules is a highly difficult operation, as it requires both the knowledge of an attacker, and the knowledge of the functionalities of all IDSEs involved in the detection process. In [59], [38], [19], we focus on the transformation process that allows to translate the description of a complex attack scenario into correlation rules and its assessment. We show that, once a human expert has provided an action tree derived from an attack tree, a fully automated transformation process can generate exhaustive correlation rules that would be tedious and error prone to enumerate by hand. The transformation relies on a detailed description of various aspects of the real execution environment (topology of the system, deployed services, etc.). Consequently, the generated correlation rules are tightly linked to the characteristics of the monitored information system. The proposed transformation process has been implemented in a prototype that generates correlation rules expressed in an attack description language called Adele. Additionally, a work has been performed to assess the approach on real environment, and to evaluate the accuracy of the rules built.

In the context of the PhD of Mouna Hkimi, we propose a approach to detect intrusions that affect the behavior of distributed applications. To determine whether an observed behavior is normal or not (occurrence of an attack), we rely on a model of normal behavior. This model has been built during an initial training phase. During this preliminary phase, the application is executed several times in a safe environment. The gathered traces (sequences of actions) are used to generate an automaton that characterizes all these acceptable behaviors. To reduce the size of the automaton and to be able to accept more general behaviors that are close to the observed traces, the automaton is transformed. These transformations may lead to introduce unacceptable behaviors. Our current work aims at identifying the possible errors tolerated by the compacted automaton.

7.1.2. Android Malware Analysis

We explore how information flows induced by a tainted application are helpful to understand how this tainted application interacts within other components inside the operating system. For that purpose, we have defined a new data structure called System Flow Graph representing in a graph how a marked data is disseminated (inside the operating system). We have shown that this data structure is helpful to understand and represent malicious behaviors [31]. Our main challenge is thus to be able to produce relevant graphs which means being able to really observe malicious executions.

For that purpose we developed GroddDroid [25] a tool dedicated to the automatic triggering of Android malware. GroddDroid makes a first static analysis of the application bytecode. During this analysis, GroddDroid identifies the suspicious parts of the bytecode and modifies the bytecode in order to exhibit an execution path that leads to these suspicious parts. The application is later reconstructed/recompiled in order to be executed. This way, GroddDroid offers a way to force the suspicious code to be executed and then observed.

7.1.3. Comparative Study of Alert Formats

In the context of the SECEF project, we conducted a comparative study of different existing alert formats [39]. We analyzed two proprietary formats, CEF (HP ArcSight) and LEEF (IBM QRADAR), as well as 4 standard formats, IDMEF (IETF), CEE (MITRE), CIM and CADF (DMTF). We proposed several metrics to compare them based on an accurate review of every fields proposed by each format. The results show that IDMEF is the most expressive and structured format. However, some fields proposed by other formats are not covered in IDMEF. We proposed some modification of the alert format to take those limitations into account.

7.1.4. Visualization

This year, research on visualization for security was oriented towards two objectives. First, as we did during the previous years, we tried to provide solution for security analysts to better analyze *a posteriori* events related to security that are happening on a system. Christopher Humphries, who was the first CIDRE Ph.D. student on this topic defended his Ph.D. Thesis *User-Centered Security Events Visualization* this December. We should also mention that we presented a prototype of our tool ELVIS during the FIC 2015 in Lille on the Pôle Cyber-Défense area.

This year, we also started research on a new topic in visualization for security. By contrast with our previous work that was dedicated to forensics, i.e. *a posteriori* analysis of security events, we started this year to study real time analysis of alerts generated by an IDS. The idea here is to allow better monitoring of what is currently happening on a system. We proposed VEGAS, a tool that allows front-line security operators to perform a first triage of the alerts to provide consistent groups of alerts to security analysts. A new Ph.D. student, Damien Crémilleux, was hired on a DGA-MI funding, to work on this topic. VEGAS was presented during the poster session of VizSec 2015 [58] that took place in Chicago, Illinois, USA on the 26th of October 2015.

7.2. Privacy

7.2.1. The Right to be Forgotten

The right to be forgotten, or to oblivion, is an aspect of privacy rights. It relates to the need for individuals to be able to leave a part of their past behind them, to change their mind about something or to take a new start in a given domain. The final report of the DAO project [53] presents an analysis of the concept from a multidisciplinary point of view, including a sociological study, a legal state of the art assorted with insights of possible evolutions, and a technical state of the art along with the proposal of a new architecture [22]. A joint technical and legal analysis of the conceptual and technical issues specific to social networks is also proposed. From the point of view of a computer scientist, the most obvious issue with the right to be forgotten is the ability to control the deletion of a piece of information once it has been disclosed and disseminated. In the general case, no guarantees can be provided, but under certain conditions it is possible to enforce remote deletion with reasonable guarantees. In general, it implies that architectural and applicative choices are made beforehand, either to allow for future decisions regarding data made available in a controlled framework, like late modifications of its access policy or the triggering of its destruction, or to plan deletion from the beginning and set a time-to-leave when disclosing the data within a particular environment, or . The approach designed in CIDRE, relying on both ephemeral publication and data degradation techniques, falls in the latter category, improving the utility for third parties (when compared to existing ephemeral publication techniques) and building a new trade-off with the users' privacy needs, by making different versions of the original data, more or less precise, available for different durations, the more detailed information being lost the quickest.

CIDRE also contributes, through the B<>com IRT, to the supervision (by Annie Blandin, professor at Télécom Bretagne, and Guillaume Piolle) of Gustav Malis's doctoral work in law in the domain of a restrictive case of the right to be forgotten. In this context, very original contributions have been made at the intersection between the two fields. In particular, a joint analysis has been proposed on the roles of legal and computing tools for the implementation of the right to be forgotten [50]. In particular, it seems that the two domains consider the issue with very different perspectives: the computer scientist almost takes for granted that he cannot rely on regulations and on "security through legality", hence the tools he designs are intended to directly empower the

user, putting him in control of his data by using preventive protection techniques. The tools may fail though, or more likely their applicability conditions may not suit all scenarios. When issues arise they may be captured by the regulatory framework, which intends to provide means for reparation and restoration. Both approaches fail to encompass all possible situations and to solve all potential issues, but they provide users and citizens with complementary tools.

The work combining computer science and law conducted in the DAO projet as well as the main conclusions of the project have also been presented in interdisciplinary colloquium by Sébastien Gams and Maryline Boizard [48], [47].

7.2.2. *Private and Secure Location-based Services*

Mobility has always been an important aspect of human activities. Nowadays problems of congestion in urban areas due to the massive usage of cars, last-minutes travel needs and progress in information and communication technologies encourage the rise of new transport modes. Among those are carpooling services, which let car owners share the empty seats of their cars with other travellers having the same travel destination. However, the way carpooling services are implemented today raises several privacy issues. In a recent paper, together with researchers from LAAS-CNRS we have proposed to use privacy enhancing technologies to improve the quality of carpooling services by specially taking in consideration privacy aspects [46].

In addition, publishing directly human mobility data raises serious privacy issues due to its inference potential, such as the (re-)identification of individuals. To address these issues and to foster the development of such applications in a privacy-preserving manner, we propose in a recent paper [26] a novel approach in which Call Detail Records (CDRs) are summarized under the form of a differentially-private Bloom filter for the purpose of privately estimating the number of mobile service users moving from one area (region) to another in a given time frame. Our sanitization method is both time and space efficient, and ensures differential privacy while solving the shortcomings of a solution recently proposed. We also report on experiments conducted using a real life CDRs dataset, which show that our method maintains a high utility while providing strong privacy.

Finally, in authentication protocols, a relay attack allows an adversary to impersonate a legitimate prover, possibly located far away from a verifier, by simply forwarding messages between these two entities. The effectiveness of such attacks has been demonstrated in practice in many environments, such as ISO 14443-compliant smartcards and car-locking mechanisms. Distance-bounding (DB) protocols, which enable the verifier to check his proximity to the prover, are a promising countermeasure against relay attacks. In such protocols, the verifier measures the time elapsed between sending a challenge and receiving the associated response of the prover to estimate their proximity. So far, distance bounding has remained mainly a theoretical concept. Indeed in practice, up to our knowledge only three ISO 14443-compliant implementations of DB protocols exist. The first two are implemented on proprietary smartcards while the last one is available on a highly-customized and dedicated hardware. In a recent paper [35], we demonstrated a proof-of-concept implementation of the Swiss-Knife DB protocol on smartphones running in RFID-emulation mode. To our best knowledge, this is the first time that such an implementation has been performed. Our experimental results are encouraging as they show that relay attacks introducing more than 1.5 ms are directly detectable (in general off-the-shelf relay attacks introduce at least 10 ms of delay). We also leverage on the full power of the ISO-DEP specification to implement the same protocol with 8-bit challenges and responses, thus reaching a better security level per execution without increasing the possibility of relay attacks. The analysis of our results leads to new promising research directions in the area of distance bounding.

7.3. Trust

Reputation mechanisms allow users to mutually evaluate their trust. This is achieved through the computation of a reputation score summarizing their past behaviors. Depending on these scores, users are free to accept or refuse to interact with each other. Existing solutions often rely on costly cryptographic tools that may lead to impractical solutions. We have proposed in [41], [40], [28] usable privacy preserving reputation mechanisms. These mechanisms are distributed and handles non-monotonic ratings. Evaluation made on our mechanism reveals it to be fully usable even with cheap on-board computers. This is a very encouraging result as it shows

that privacy does not impede utility and accuracy. This has been achieved by combining distributed algorithms and cryptographic schemes. Our mechanism is independent of the reputation model, that is, our system can integrate any reputation model, preferably one using both positive and negative ratings.

In a mobile ad hoc network we have also considered the problem of designing a reputation system that allows to update and to propagate the computed reputation scores while tolerating Byzantine failures [42]. Each time a correct node uses directly a service, it can determine by itself the quality of service currently provided. This fresh and valid rating information is broadcast immediately to all its current neighbors. Then, while the mobile node moves, it can receive from other nodes other recommendations also related to the same service. Thus it updates continuously its own opinion. Meanwhile it continues to broadcast this updated information. The freshness and the validity of the received/sent information become questionable. We propose a protocol that allows a node to ignore a second hand information when this information is not fresh or not valid. In particular, fake values provided by Byzantine nodes are eliminated when they are not consistent with those gathered from correct nodes. When the quality of service stabilizes, the correct nodes are supposed to provide quite similar recommendations. In this case, we demonstrate that the proposed protocol ensures convergence to a range of possible reputation scores if a necessary condition is satisfied by the mobile nodes. Simulations are conducted in random mobility scenarios. The results show that our algorithm has a better performance than typical methods proposed in previous works.

7.4. Other Topics Related to Security or Distributed Computing

7.4.1. Detection of distributed denial of service attacks

A Denial of Service (DoS) attack tries to progressively take down an Internet resource by flooding this resource with more requests than it is capable to handle. A Distributed Denial of Service (DDoS) attack is a DoS attack triggered by thousands of machines that have been infected by a malicious software, with as immediate consequence the total shut down of targeted web resources (*e.g.*, e-commerce websites). A solution to detect and to mitigate DDoS attacks is to monitor network traffic at routers and to look for highly frequent signatures that might suggest ongoing attacks. A recent strategy followed by the attackers is to hide their massive flow of requests over a multitude of routes, so that locally, these flows do not appear as frequent, while globally they represent a significant portion of the network traffic. The term “iceberg” has been recently introduced to describe such an attack as only a very small part of the iceberg can be observed from each single router. The approach adopted to defend against such new attacks is to rely on multiple routers that locally monitor their network traffic, and upon detection of potential icebergs, inform a monitoring server that aggregates all the monitored information to accurately detect icebergs [29]. Now to prevent the server from being overloaded by all the monitored information, routers continuously keep track of the c (among n) most recent high flows (modeled as items) prior to sending them to the server, and throw away all the items that appear with a small probability p_i , and such that the sum of these small probabilities is modeled by probability p_0 . Parameter c is dimensioned so that the frequency at which all the routers send their c last frequent items is low enough to enable the server to aggregate all of them and to trigger a DDoS alarm when needed. This amounts to compute the time needed to collect c distinct items among n frequent ones. A thorough analysis of the time needed to collect c distinct items appears in [16], [15].

7.4.2. Metrics Estimation on Very Large Data Streams

Huge data flows have become very common in the last decade. This has motivated the design of online algorithms that allow the accurate estimation of statistics on very large data flows. A rich body of algorithms and techniques have been proposed for the past several years to efficiently compute statistics on massive data streams. In particular, estimating the number of times data items recur in data streams in real time enables, for example, the detection of worms and denial of service attacks in intrusion detection services, or the traffic monitoring in cloud computing applications. Two main approaches exist to monitor in real time massive data streams. The first one consists in regularly sampling the input streams so that only a limited amount of data items is locally kept. This allows to exactly compute functions on these samples. However, accuracy of this computation with respect to the stream in its entirety fully depends on the volume of data items that has been

sampled and their order in the stream. In contrast, the streaming approach consists in scanning each piece of data of the input stream on the fly, and in locally keeping only compact synopses or *sketches* that contain the most important information about these data. This approach enables us to derive some data streams statistics with guaranteed error bounds without making any assumptions on the order in which data items are received at nodes. Sketches highly rely on the properties of hashing functions to extract statistics from them. Sketches vary according to the number of hash functions they use, and the type of operations they use to extract statistics. The *Count-Min sketch* algorithm proposed by Cormode and Muthukrishnan in 2005 so far predominates all the other ones in terms of space and time needed to guarantee an additive ϵ -accuracy on the estimation of item frequencies. Briefly, this technique performs t random projections of the set of items of the input stream into a much smaller co-domain of size k , with $k = \lceil e/\epsilon \rceil$ and $t = \lceil \log(1/\delta) \rceil$ in which $0 < \epsilon, \delta < 1$. The user defined parameters ϵ and δ represent respectively the accuracy of the approximation, and the probability with which the accuracy holds. However, because k is typically much smaller than the total number of distinct items in the input stream, hash collisions do occur. This affects the estimation of item frequency when the size of the stream is large. In this work, we have proposed an alternative approach to reduce the impact of collisions on the estimation of item frequency. The intuition of our idea is that by keeping track of the most frequent items of the stream, and by removing their weight from the one of the items with which these frequent items collide, the over-estimation of non frequent items is drastically decreased [21].

We have also proposed a metric, called codeviation, that allows to evaluate the correlation between distributed streams [27]. This metric is inspired from classical metric in statistics and probability theory, and as such allows us to understand how observed quantities change together, and in which proportion. We then propose to estimate the codeviation in the data stream model. In this model, functions are estimated on a huge sequence of data items, in an online fashion, and with a very small amount of memory with respect to both the size of the input stream and the values domain from which data items are drawn. We give upper and lower bounds on the quality of the codeviation, and provide both local and distributed algorithms that additively approximates the codeviation among n data streams by using a sublinear number of bits of space in the size of the domain value from which data items are drawn, and the maximal stream length. To the best of our knowledge, such a metric has never been proposed so far.

7.4.3. Stream Processing Systems

Stream processing systems are today gaining momentum as a tool to perform analytics on continuous data streams. Their ability to produce analysis results with sub-second latencies, coupled with their scalability, makes them the preferred choice for many big data companies.

A stream processing application is commonly modeled as a direct acyclic graph where data operators, represented by nodes, are interconnected by streams of tuples containing data to be analyzed, the directed edges. Scalability is usually attained at the deployment phase where each data operator can be parallelized using multiple instances, each of which will handle a subset of the tuples conveyed by the operator's ingoing stream. Balancing the load among the instances of a parallel operator is important as it yields to better resource utilization and thus larger throughputs and reduced tuple processing latencies. We have proposed a new key grouping technique targeted toward applications working on input streams characterized by a skewed value distribution [44]. Our solution is based on the observation that when the values used to perform the grouping have skewed frequencies, e.g. they can be approximated with a Zipfian distribution, the few most frequent values (the *heavy hitters*) drive the load distribution, while the remaining largest fraction of the values (the *sparse items*) appear so rarely in the stream that the relative impact of each of them on the global load balance is negligible. We have shown, through a theoretical analysis, that our solution provides on average near-optimal mappings using sub-linear space in the number of tuples read from the input stream in the learning phase and the support (value domain) of the tuples. In particular this analysis presents new results regarding the expected error made on the estimation of the frequency of heavy hitters.

7.4.4. Randomized Message-Passing Test-and-Set

In [30], we have presented a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes,

returns yes to a unique process and returns no to all the others. Recently many advances in implementing Test&Set objects have been achieved, however all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p \leq n$ of processes where n is the total number of processes in the system. It has an expected individual step complexity in $O(\log p)$ against an oblivious adversary, and an expected individual message complexity in $O(n)$. The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

7.4.5. Population Protocol Model

The population protocol model, introduced by Angluin et his colleagues in 2006, provides theoretical foundations for analyzing global properties emerging from pairwise interactions among a large number of anonymous agents. In the population protocol model, agents are modeled as identical and deterministic finite state machines, *i.e.* each agent can be in a finite number of states while waiting to execute a transition. When two agents interact, they communicate their local state, and can move from one state to another according to a joint transition function. The patterns of interaction are unpredictable, however they must be fair, in the sense that any interaction that should possibly appear cannot be avoided forever. The ultimate goal of population protocols is for all the agents to converge to a correct value independently of the interaction pattern. Examples of systems whose behavior can be modeled by population protocols range from molecule interactions of a chemical process to sensor networks in which agents, which are small devices embedded on animals, interact each time two animals are in the same radio range.

In this work, we focus on an quite important related question. Namely, is there a population protocol that exactly counts the difference κ between the number of agents that initially set their state to A and the one that initially set it to B , and can it be solved in an efficient way, that is with the guarantee that each agent should converge to the exact value of κ after having triggered a sub-linear number of interactions in the size of the system [43].

We answer this question by the affirmative by presenting a $O(n^{3/2})$ -state population protocol that allows each agent to converge to the exact solution by interacting no more than $O(\log n)$ times. The proposed protocol is very simple (as is true for most known population protocols), but is general enough to be used to solve different types of tasks.

COAST Project-Team

5. New Results

5.1. Probabilistic Partial Orderings

Participants: Jordi Martori Adrian, Pascal Urso.

Ordering events in a distributed system fundamentally consists in delaying event delivery. Partial ordering, such as FIFO and causal order, has many usage in practical distributed and collaborative systems and can be obtained in arbitrarily large and dynamic networks. However, partial orderings imply that messages cannot be sent and delivered as soon as produced.

In [14], we study the latency induced by such partial orderings. We obtain a probabilistic measure of the moment a message can be delivered according the different characteristics of the distributed system. Having such a measure helps to understand the systems behaviour and to design new protocols. For instance, our measure allows us to parametrize a naive, albeit efficient, fault-tolerant causal delivery mechanism. We experimentally validate our approach using Internet-scale production distribution latency including faults.

5.2. Effect of Delay on Group Performance

Participants: François Charoy, Claudia-Lavinia Ignat [contact], Gérald Oster.

We continued our work on studying the effect of delay in real-time collaborative editing. Delays exist between the execution of one user's modification and the visibility of this modification to the other users. Such delays are in part fundamental to the network, as well as arising from the consistency maintenance algorithms and underlying architecture of collaborative editors. Existing quantitative research on collaborative document editing does not examine either concern for delay or the efficacy of compensatory strategies.

In [12] we studied a collaborative note taking task where we introduced simulated delay. The study was done with 20 groups of 4 users which were asked to listen to a short interview and take notes. We found out a general effect of delay on performance related to the ability to manage redundancy and errors across the document. We interpret this finding as a compromised ability to maintain awareness of team member activity, and a reversion to independent work. Measures of common ground in accompanying chat indicate that groups with less experienced team members attempt to compensate for the effect of delay. In contrast, more experienced groups do not adjust their communication in response to delay, and their performance remains sensitive to the delay manipulation. Results of this study support our team assertion that delay associated with conventional consistency maintenance algorithms will impede group performance. Therefore, these results promote the use of novel algorithms such as CRDTs and motivate the pursuance of research and development on these approaches.

5.3. A CRDT Supporting Selective Undo for Collaborative Text Editing

Participants: Luc André, Claudia-Lavinia Ignat [contact].

Selective undo is an important feature in collaborative editors. With selective undo, a user can undo an earlier operation, regardless of when and where the operation was generated. Current systems that support selective undo are subject to two main limitations. Firstly, they only support undo of operations on atomic objects (e.g. characters or un-breakable lines). In the case of string-wise operations such as copy-paste, find-replace or select-delete, users can typically only undo earlier operations character by character. Secondly, selective undo may lead to undesirable effects. For example, a user first inserts a misspelled word and then makes a correction. The correction depends on the first insertion of the word. It is undesirable to undo the insertion alone and leave the correction behind as a groundless modification. In [15] we proposed a novel consistency maintenance approach relying on a layered commutative replicated data type (CRDT) that supports selective undo of string-wise operations in collaborative editing. This is the first work that manages undesirable effects of undo. Our performance study shows that it provides sufficient responsiveness to the end users.

5.4. A Trust-Based Formal Delegation Framework for Enterprise Social Networks

Participants: Ahmed Bouchami, Olivier Perrin [contact].

Collaborative environments raise major challenges to secure them. These challenges increase when it comes to the domain of Enterprise Social Networks (ESNs) as ESNs aim to incorporate the social technologies in an organization setup while asserting greater control of information security. In this context, the security challenges have taken a new shape as an ESN may not be limited to the boundaries of a single organization and users from different organizations can collaborate in a common federated environment.

We address the problem of the authorization's delegation in federated collaborative environments like ESNs with an approach based on event-calculus, a temporal logic programming formalism. While the traditional approaches are either user-centric or organization-centric, the approach bridges the gap between these two views and the proposed framework enhances the delegation scheme. We have proposed a behavior monitoring mechanism, that permits to assess principals' trust level within the federated collaborative environment [10].

5.5. Risk Management in the Cloud. Application to Business Process Deployment

Participants: Claude Godart [contact], Elio Goettelmann.

The lack of trust in cloud organizations is often seen as braking forces to SaaS developments. This work proposes an approach which supports a trust model and a business process model in order to allow the orchestration of trusted business process components in the cloud.

The contribution is threefold and consists in a method, a model and a framework. The method categorizes techniques to transform an existing business process into a risk-aware process model that takes into account security risks related to cloud environments. These techniques are partially described in the form of constraints to automatically support process transformation. The model formalizes the relations and the responsibilities between the different actors of the cloud. This allows to identify the different information required to assess and quantify security risks in cloud environments.

The framework is a comprehensive approach that decomposes a business process into fragments that can automatically be deployed on multiple clouds. The framework also integrates a selection algorithm that combines the security information of cloud offers and of the process with other quality of service criteria to generate an optimized configuration. It is implemented in a tool to assess cloud providers.

Elio Goettelmann has defended his PhD thesis entitled "Risk-aware Business Process Modeling and Trusted Deployment in the Cloud" on October 2015 [1] based on this result. This framework has been combined to an access control model for strengthening access controls in the context of a collaborative federation of components [9].

5.6. Secure Business Process Deployment in SaaS Contexts

Participants: Amina Ahmed Nacer, Claude Godart [contact], Elio Goettelmann, Samir Youcef.

Business process (BP) stakeholders want the benefits of the cloud, but they are also reluctant to expose their BP models which express the know-how of their companies. To prevent such a know-how exposure, we are developing a design-time approach for obfuscating a BP model by splitting its model into a collaboration of BP fragments semantically equivalent to the initial BP. This breaking down renders the discovery of the deep content of a critical fragment or of the whole process semantics, by cloud providers much harder when these fragments are deployed in a multi-cloud context. While existing contributions on this topic remain at the level of principles, we propose an algorithm supporting such a BP model transformation [11]. To validate this approach, we are developing a new metric of obfuscation. Complementary to obfuscation, we are developing techniques to reuse, at design time, business process fragments from the cloud, but with limited security risks [8].

5.7. Web Services Selection with QoS

Participants: Amina Ahmed Nacer, Kahina Bessai, Claude Godart [contact], Samir Youcef.

The development of the web technologies and the increase of available services raise the issue of the selection of the most appropriate service among a set of candidate web services. First of all, the services offering a given functionality are discovered. Then, the service selection process assists users in choosing the services that better meets their preferences. These preferences are generally expressed as potentially objective functions often conflicting.

Most of existing works trying to select the best web services are based either on a single evaluation criterion or, at best, on the use of an aggregation function like weighted sum of several quantitative evaluation criteria, or the use of the Pareto optimality notion.

In this work, we address some shortcomings of existing approaches by introducing a new optimality notion based on two tests: (i) concordance and (ii) discordance tests. It presents an efficient algorithm to select only the best services using the introduced optimality notion. Moreover, the proposed algorithm exhibits encouraging results as supported by a series of experiments [7].

COATI Project-Team

7. New Results

7.1. Network Design and Management

Participants: Jean-Claude Bermond, Christelle Caillouet, David Coudert, Frédéric Giroire, Frédéric Havet, Nicolas Huin, Alvinice Kodjo, Fatima Zahra Moataz, Joanna Moulhierac, Nicolas Nisse, Stéphane Pérennes.

7.1.1. Wireless Networks

7.1.1.1. Dimensioning Microwave Wireless Networks

In [47], we aim at dimensioning fixed broadband microwave wireless networks under unreliable channel conditions. As the transport capacity of microwave links is prone to variations due to, e.g., weather conditions, such a dimensioning requires special attention. It can be formulated as the determination of the minimum cost bandwidth assignment of the links in the network for which traffic requirements can be met with high probability, while taking into account that transport link capacities vary depending on channel conditions. The proposed optimization model represents a major step forward since we consider dynamic routing. Experimental results show that the resulting solutions can save up to 45% of the bandwidth cost compared to the case where a bandwidth over-provisioning policy is uniformly applied to all links in the network planning. Comparisons with previous work also show that we can solve much larger instances in significantly shorter computing times, with a comparable level of reliability.

7.1.1.2. Data Gathering and Personalized Broadcasting in Radio Grids with Interference

In the gathering problem, a particular node in a graph, the base station, aims at receiving messages from some nodes in the graph. At each step, a node can send one message to one of its neighbors (such an action is called a call). However, a node cannot send and receive a message during the same step. Moreover, the communication is subject to interference constraints, more precisely, two calls interfere in a step, if one sender is at distance at most dI from the other receiver. Given a graph with a base station and a set of nodes having some messages, the goal of the gathering problem is to compute a schedule of calls for the base station to receive all messages as fast as possible, i.e., minimizing the number of steps (called makespan). The gathering problem is equivalent to the personalized broadcasting problem where the base station has to send messages to some nodes in the graph, with same transmission constraints. In [24], we focus on the gathering and personalized broadcasting problem in grids. Moreover, we consider the non-buffering model: when a node receives a message at some step, it must transmit it during the next step. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we present linear (in the number of messages) algorithms that compute schedules for gathering with $dI \in \{0, 1, 2\}$. In particular, we present an algorithm that achieves the optimal makespan up to an additive constant 2 when $dI = 0$. If no messages are “close” to the axes (the base station being the origin), our algorithms achieve the optimal makespan up to an additive constant 1 when $dI = 0$, 4 when $dI = 2$, and 3 when both $dI = 1$ and the base station is in a corner. Note that, the approximation algorithms that we present also provide approximation up to a ratio 2 for the gathering with buffering. All our results are proved in terms of personalized broadcasting.

7.1.2. Elastic Optical Networks

7.1.2.1. On Spectrum Assignment in Elastic Optical Tree-Networks

To face the explosion of the Internet traffic, a new generation of optical networks is being developed; the Elastic optical Networks (EONs). The aim with EONs is to use the optical spectrum efficiently and flexibly. The benefit of the flexibility is, however, accompanied by more difficulty in the resource allocation problems. In [54], [51], [14], we study the problem of Spectrum Allocation in Elastic Optical Tree-Networks. In trees, even though the routing is fixed, the spectrum allocation is NP-hard. We survey the complexity and approximability results that have been established for the SA in trees and prove new results for stars and binary trees.

7.1.3. Fault Tolerance

7.1.3.1. Shared Risk Link Group

The notion of Shared Risk Link Groups (SRLG) captures survivability issues when a set of links of a network may fail simultaneously. The theory of survivable network design relies on basic combinatorial objects that are rather easy to compute in the classical graph models: shortest paths, minimum cuts, or pairs of disjoint paths. In the SRLG context, the optimization criterion for these objects is no longer the number of edges they use, but the number of SRLGs involved. Unfortunately, computing these combinatorial objects is NP-hard and hard to approximate with this objective in general. Nevertheless some objects can be computed in polynomial time when the SRLGs satisfy certain structural properties of locality which correspond to practical ones, namely the star property (all links affected by a given SRLG are incident to a unique node) and the span 1 property (the links affected by a given SRLG form a connected component of the network). The star property is defined in a multi-colored model where a link can be affected by several SRLGs while the span property is defined only in a mono-colored model where a link can be affected by at most one SRLG. In [59], we extend these notions to characterize new cases in which these optimization problems can be solved in polynomial time. We also investigate the computational impact of the transformation from the multi-colored model to the mono-colored one. Experimental results are presented to validate the proposed algorithms and principles.

In [22], we investigate the k -diverse routing problem which is to find a set of k pairwise SRLG-disjoint paths between a given pair of end nodes of the network. This problem has been proven NP-complete in general and some polynomial instances have been characterized. We consider more specifically the case where the SRLGs are localized and satisfy the star property. We first provide counterexamples to the polynomial time algorithm proposed by X. Luo and B. Wang (DRCN'05) for computing a pair of SRLG-disjoint paths in networks with SRLGs satisfying the star property, and then prove that this problem is in fact NP-complete. We then characterize instances that can be solved in polynomial time or are fixed parameter tractable, in particular when the number of SRLGs is constant, the maximum degree of the vertices is at most 4, and when the network is a directed acyclic graph. Finally we consider the problem of finding the maximum number of SRLG-disjoint paths in networks with SRLGs satisfying the star property. We prove that this problem is NP-hard to approximate within $O(|V|^{1-\varepsilon})$ for any $0 < \varepsilon < 1$, where V is the set of nodes in the network. Then, we provide exact and approximation algorithms for relevant subcases.

7.1.3.2. Design of Fault-tolerant On-board Networks with Variable Switch Sizes

In [29], we focus on designing networks that are capable, in the presence of faulty output ports, of rerouting input signals to operational output ports. Since the components of a satellite cannot be repaired, redundant amplifiers are added, and the interconnection network satisfies the following fault tolerance property: the network connects the set of input ports with the set of output ports, and for any set of at most k output port failures, there exists a set of edge-disjoint paths connecting the input ports to the operational output ports. Since each switching device is expensive, these interconnection networks are constructed using the fewest possible switches, or at least a number of switches close to the minimum value. The networks are controlled centrally from Earth. Each time an amplifier in use develops a fault, the controller sends messages to the switches to change their settings, so as to ensure that the inputs remain connected to functioning amplifiers.

Current switches have four ports. Obviously, the larger the number of ports, the more expensive will be the switches, but then fewer will be required. So the cost of such a network involves a trade-off between the total number of switches and their unit cost. In order to determine the minimum-cost network, we give some bounds on the minimum number $\mathcal{N}(n, k, r)$ of $2r$ -port switches in interconnection networks with n inputs and $n + k$ outputs.

We first show $\mathcal{N}(n, k, r) \leq \left\lceil \frac{k+2}{2r-2} \right\rceil \left\lceil \frac{n}{2} \right\rceil$. When $r \geq k/2$, we prove a better upper bound: $\mathcal{N}(n, k, r) \leq \frac{r-2+k/2}{r^2-2r+k/2}n + O(1)$. Next, we establish some lower bounds. We show that if $k \geq r$, then $\mathcal{N}(n, k, r) \geq \frac{3n+k}{2r}$. We improve this bound when $k \geq 2r$: $\mathcal{N}(n, k, r) \geq \frac{3n+2k/3-r/2}{2r-2+\lfloor \frac{k}{r} \rfloor}$. Finally, we determine $\mathcal{N}(n, k, r)$ up to additive constants for $k \leq 6$.

7.1.4. Reducing Networks' Energy Consumption

Due to the increasing impact of ICT (Information and Communication Technology) on power consumption and worldwide gas emissions, energy efficient ways to design and operate backbone networks are becoming a new concern for network operators. Recently, energy-aware routing (EAR) has gained an increasing popularity in the networking research community. The idea is that traffic demands are redirected over a subset of the network links, allowing other links to sleep to save energy. We studied variant of this problems.

7.1.4.1. Robust Energy-aware Routing with Redundancy Elimination

In [31], we propose GreenRE – a new EAR model with the support of data redundancy elimination (RE). This technique, enabled within routers, can virtually increase the capacity of network links. Based on real experiments on a Orange Labs platform, we show that performing RE increases the energy consumption for routers. Therefore, it is important to determine which routers should enable RE and which links to put into sleep mode so that the power consumption of the network is minimized. We model the problem as a Mixed Integer Linear Program and propose greedy heuristic algorithms for large networks. Simulations on several network topologies show that the GreenRE model can gain further 37% of energy savings compared to the classical EAR model. In [27], we introduce an extended model of the classical multi-commodity flow problem with compressible flows which is also robust with fluctuation of traffic demand and compression rate. An heuristic built on this model allows for 16-28% extra energy saving.

7.1.4.2. Optimizing IGP Link Weights for Energy-efficiency in Multi-period Traffic Matrices

To guarantee QoS while implementing EAR, all traffic demands should be routed without violating capacity constraints and the network should keep its connectivity. From the perspective of traffic engineering, we argue that stability in routing configuration also plays an important role in QoS. In details, frequent changes in network configuration (link weights, slept and activated links) to adapt with traffic fluctuation in daily time cause network oscillations. In [35], we propose a novel optimization method to adjust the link weights of Open Shortest Path First (OSPF) protocol while limiting the changes in network configurations when multi-period traffic matrices are considered.

7.1.4.3. Energy Efficient Content Distribution

Recently, there is a trend to introduce content caches as an inherent capacity of network equipment, with the objective of improving the efficiency of content distribution and reducing the network congestion. In [18], we study the impact of using in-network caches and content delivery network (CDN) cooperation on an EAR. We formulate this problem as Energy Efficient Content Distribution, we propose an integer linear program (ILP) and a heuristic algorithm to solve it. The objective of this problem is to find a feasible routing, so that the total energy consumption of the network is minimized while the constraints given by the demands and the link capacity are satisfied. We exhibit for which the range of parameters (size of caches, popularity of content, demand intensity, etc.) it is useful to use caches. Experimental results show that by placing a cache on each backbone router to store the most popular content, along with well choosing the best content provider server for each demand to a CDN, we can save about 20% of power in average of all the backbone networks considered.

7.1.5. Routing Theory and Forwarding Index

Motivated by finding the best set of links that should be on for energy efficiency, we study the problem of determining the minimum forwarding index of a graph. The (edge) forwarding index of a graph is the minimum, over all possible routings of all the demands, of the maximum load of an edge. This metric is of a great interest since it captures the notion of global congestion in a precise way: the lesser the forwarding-index, the lesser the congestion. This parameter has been studied for different graph classes in the literature. In [42], we determine, for different numbers of edges, the best spanning graphs of a square grid, namely those with a low forwarding index. In [61], [43], we study the following design question: Given a number e of edges and a number n of vertices, what is the least congested graph that we can construct? and what forwarding-index can we achieve? We answer here these questions for different families of graphs: general graphs, graphs with bounded degree, sparse graphs with a small number of edges by providing constructions, most of them

asymptotically optimal. Doing so, we partially answer the practical problem that initially motivated our work: If an operator wants to power only e links of its network, in order to reduce the energy consumption (or wiring cost) of its networks, what should be those links and what performance can be expected?

On the complexity of equal shortest path routing.

Additionally, we studied the complexity of configuring the OSPF-ECMP (for Open Shortest Path First-Equal Cost Multiple Path) protocol. In [32], we show that the problem of maximizing even a single commodity flow for the OSPF-ECMP protocol cannot be approximated within any constant factor ratio. Besides this main theorem, we derive some positive results which include polynomial-time approximations and an exponential-time exact algorithm.

7.1.6. Routing in Software Defined Networks (SDN)

Software Defined Networking (SDN) is gaining momentum with the support of major manufacturers. While it brings flexibility in the management of flows within the data center fabric, this flexibility comes at the cost of smaller routing table capacities. In [50], we investigate compression techniques to reduce the forwarding information base (FIB) of SDN switches. We validate our algorithm, called MINNIE, on a real testbed able to emulate a 20 switches fat tree architecture. We demonstrate that even with a small number of clients, the limit in terms of number of rules is reached if no compression is performed, increasing the delay of all new incoming flows. MINNIE, on the other hand, reduces drastically the number of rules that need to be stored with a limited impact on the packet loss rate. We also evaluate the actual switching and reconfiguration times and the delay introduced by the communications with the controller. In parallel, we considered the algorithmic problem of compressing bidimensional routings table with priorities on the rules. We carry out in [40] a study of the problem complexity, providing results of NP-completeness, of Fixed-Parameter Tractability and approximation algorithms. In [44], we then propose green routing schemes performing simultaneously the selection of the routes, the compression of the routing tables, and decide to put in sleep mode unused links. These algorithms are tested on networks from the SNDLib library.

7.1.7. Video Streaming

7.1.7.1. Study of Repair Protocols for Live Video Streaming Distributed Systems

In [41], we study distributed systems for live video streaming. These systems can be of two types: structured and un-structured. In an unstructured system, the diffusion is done opportunistically. The advantage is that it handles churn, that is the arrival and departure of users, which is very high in live streaming systems, in a smooth way. On the opposite, in a structured system, the diffusion of the video is done using explicit diffusion trees. The advantage is that the diffusion is very efficient, but the structure is broken by the churn. In this paper, we propose simple distributed repair protocols to maintain, under churn, the diffusion tree of a structured streaming system. We study these protocols using formal analysis and simulation. In particular, we provide an estimation of the system metrics, bandwidth usage, delay, or number of interruptions of the streaming. Our work shows that structured streaming systems can be efficient and resistant to churn.

7.2. Graph Algorithms

Participants: Nathann Cohen, David Coudert, Frédéric Giroire, Fatima Zahra Moataz, Benjamin Momège, Nicolas Nisse, Stéphane Pérennes.

COATI is also interested in the algorithmic aspects of Graph Theory. In general we try to find the most efficient algorithms to solve various problems of Graph Theory and telecommunication networks. We use graph theory to model various network problems. We study their complexity and then we investigate the structural properties of graphs that make these problems hard or easy. In particular, we try to find the most efficient algorithms to solve the problems, sometimes focusing on specific graph classes from which the problems are polynomial-time solvable. Many results introduced here are presented in detail in the PhD thesis of F. Z. Moataz [14].

7.2.1. Graph Hyperbolicity

The Gromov hyperbolicity is an important parameter for analyzing complex networks which expresses how the metric structure of a network looks like a tree (the smaller gap the better). It has recently been used to provide bounds on the expected stretch of greedy-routing algorithms in Internet-like graphs, and for various applications in network security, computational biology, the analysis of graph algorithms, and the classification of complex networks.

7.2.1.1. Exact Algorithms for Computing the Gromov Hyperbolicity

The best known theoretical algorithm computing this parameter runs in $O(n^{3.69})$ time, which is prohibitive for large-scale graphs. In [26], we propose an algorithm for determining the hyperbolicity of graphs with tens of thousands of nodes. Its running time depends on the distribution of distances and on the actual value of the hyperbolicity. Although its worst case runtime is $O(n^4)$, it is in practice much faster than previous proposals as observed in our experimentations on benchmark instances. We also propose a heuristic algorithm that can be used on graphs with millions of nodes.

In [37], we provide a more efficient algorithm: although its worst-case complexity remains in $O(n^4)$, in practice it is much faster, allowing, for the first time, the computation of the hyperbolicity of graphs with up to 200,000 nodes. We experimentally show that our new algorithm drastically outperforms the best previously available algorithms, by analyzing a big dataset of real-world networks. We have also used the new algorithm to compute the hyperbolicity of random graphs generated with the Erdős-Renyi model, the Chung-Lu model, and the Configuration Model.

7.2.1.2. Hyperbolicity of Particular Graph Classes

Topologies for data center networks have been proposed in the literature through various graph classes and operations. A common trait to most existing designs is that they enhance the symmetric properties of the underlying graphs. Indeed, symmetry is a desirable property for interconnection networks because it minimizes congestion problems and it allows each entity to run the same routing protocol. However, despite sharing similarities these topologies all come with their own routing protocol. Recently, generic routing schemes have been introduced which can be implemented for any interconnection networks. The performances of such universal routing schemes are intimately related to the hyperbolicity of the topology. Motivated by the good performances in practice of these new routing schemes, we propose in [56] the first general study of the hyperbolicity of data center interconnection networks. Our findings are disappointingly negative: we prove that the hyperbolicity of most data center topologies scales linearly with their diameter, that it the worst-case possible for hyperbolicity. To obtain these results, we introduce original connection between hyperbolicity and the properties of the endomorphism monoid of a graph. In particular, our results extend to all vertex and edge-transitive graphs. Additional results are obtained for de Bruijn and Kautz graphs, grid-like graphs and networks from the so-called Cayley model.

In [57], we investigate more specifically on the hyperbolicity of bipartite graphs. More precisely, given a bipartite graph $B = (V_0 \cup V_1, E)$ we prove it is enough to consider any one side V_i of the bipartition of B to obtain a close approximate of its hyperbolicity $\delta(B)$ — up to an additive constant 2. We obtain from this result the sharp bounds $\delta(G) - 1 \leq \delta(L(G)) \leq \delta(G) + 1$ and $\delta(G) - 1 \leq \delta(K(G)) \leq \delta(G) + 1$ for every graph G , with $L(G)$ and $K(G)$ being respectively the line graph and the clique graph of G . Finally, promising extensions of our techniques to a broader class of intersection graphs are discussed and illustrated with the case of the biclique graph $BK(G)$, for which we prove $(\delta(G) - 3)/2 \leq \delta(BK(G)) \leq (\delta(G) + 3)/2$.

7.2.2. Tree-decompositions

We study the computational complexity of different variants of tree-decompositions. We also study their relationship with various pursuit-evasion games.

7.2.2.1. Diameter of Minimal Separators in Graphs (structure vs metric in graphs)

In [39], we establish general relationships between the topological properties of graphs and their metric properties. For this purpose, we upper-bound the diameter of the *minimal separators* in any graph by a function of their sizes. More precisely, we prove that, in any graph G , the diameter of any minimal separator S in G is

at most $\lfloor \frac{\ell(G)}{2} \rfloor \cdot (|S| - 1)$ where $\ell(G)$ is the maximum length of an isometric cycle in G . We refine this bound in the case of graphs admitting a *distance preserving ordering* for which we prove that any minimal separator S has diameter at most $2(|S| - 1)$. Our proofs are mainly based on the property that the minimal separators in a graph G are connected in some power of G . Our result easily implies that the *treelength* $tl(G)$ of any graph G is at most $\lfloor \frac{\ell(G)}{2} \rfloor$ times its *treewidth* $tw(G)$. In addition, we prove that, for any graph G that excludes an *apex graph* H as a minor, $tw(G) \leq c_H \cdot tl(G)$ for some constant c_H only depending on H . We refine this constant when G has bounded genus. As a consequence, we obtain a very simple $O(\ell(G))$ -approximation algorithm for computing the treewidth of n -node m -edge graphs that exclude an apex graph as a minor in $O(nm)$ -time.

7.2.2.2. Minimum Size Tree-decompositions

Tree-decompositions are the cornerstone of many dynamic programming algorithms for solving graph problems. Since the complexity of such algorithms generally depends exponentially on the width (size of the bags) of the decomposition, much work has been devoted to compute tree-decompositions with small width. However, practical algorithms computing tree-decompositions only exist for graphs with treewidth less than 4. In such graphs, the time-complexity of dynamic programming algorithms is dominated by the size (number of bags) of the tree-decompositions. It is then interesting to minimize the size of the tree-decompositions. In [48], [14], we consider the problem of computing a tree-decomposition of a graph with width at most k and minimum size. We prove that the problem is NP-complete for any fixed $k \geq 4$ and polynomial for $k \leq 2$; for $k = 3$, we show that it is polynomial in the class of trees and 2-connected outerplanar graphs.

7.2.2.3. Non-deterministic Graph Searching in Trees

Non-deterministic graph searching was introduced by Fomin et al. to provide a unified approach for pathwidth, treewidth, and their interpretations in terms of graph searching games. Given $q \geq 0$, the q -limited search number, $s_q(G)$, of a graph G is the smallest number of searchers required to capture an invisible fugitive in G , when the searchers are allowed to know the position of the fugitive at most q times. The search parameter $s_0(G)$ corresponds to the pathwidth of a graph G , and $s_\infty(G)$ to its treewidth. Determining $s_q(G)$ is NP-complete for any fixed $q \geq 0$ in general graphs and $s_0(T)$ can be computed in linear time in trees, however the complexity of the problem on trees has been unknown for any $q > 0$. We introduce in [16] a new variant of graph searching called restricted non-deterministic. The corresponding parameter is denoted by rs_q and is shown to be equal to the non-deterministic graph searching parameter s_q for $q = 0, 1$, and at most twice s_q for any $q \geq 2$ (for any graph G). Our main result is a polynomial time algorithm that computes $rs_q(T)$ for any tree T and any $q \geq 0$. This provides a 2-approximation of $s_q(T)$ for any tree T , and shows that the decision problem associated to s_1 is polynomial in the class of trees. Our proofs are based on a new decomposition technique for trees which might be of independent interest.

7.2.2.4. k -Chordal Graphs: from Cops and Robber to Compact Routing via Treewidth

Cops and robber games, introduced by Winkler and Nowakowski and independently defined by Quilliot, concern a team of cops that must capture a robber moving in a graph. We consider in [34] the class of k -chordal graphs, i.e., graphs with no induced (chordless) cycle of length greater than k , $k \geq 3$. We prove that $k-1$ cops are always sufficient to capture a robber in k -chordal graphs. This leads us to our main result, a new structural decomposition for a graph class including k -chordal graphs. We present a polynomial-time algorithm that, given a graph G and $k \geq 3$, either returns an induced cycle larger than k in G , or computes a tree-decomposition of G , each bag of which contains a dominating path with at most $k-1$ vertices. This allows us to prove that any k -chordal graph with maximum degree Δ has treewidth at most $(k-1)(\Delta-1) + 2$, improving the $O(\Delta(\Delta-1)k-3)$ bound of Bodlaender and Thilikos (1997). Moreover, any graph admitting such a tree-decomposition has small hyperbolicity. As an application, for any n -vertex graph admitting such a tree-decomposition, we propose a compact routing scheme using routing tables, addresses and headers of size $O(k \log \Delta + \log n)$ bits and achieving an additive stretch of $O(k \log \Delta)$. As far as we know, this is the first routing scheme with $O(k \log \Delta + \log n)$ -routing tables and small additive stretch for k -chordal graphs.

7.2.2.5. Connected Surveillance Game

The surveillance game [68] models the problem of web-page prefetching as a pursuit evasion game played on a graph. This two-player game is played turn-by-turn. The first player, called the observer, can mark a fixed

amount of vertices at each turn. The second one controls a surfer that stands at vertices of the graph and can slide along edges. The surfer starts at some initially marked vertex of the graph, its objective is to reach an unmarked node before all nodes of the graph are marked. The surveillance number $sn(G)$ of a graph G is the minimum amount of nodes that the observer has to mark at each turn ensuring it wins against any surfer in G . Fomin et al. also defined the connected surveillance game where the observer must ensure that marked nodes always induce a connected subgraph. They ask what is the cost of connectivity, i.e., is there a constant $c > 0$ such that the ratio between the connected surveillance number $csn(G)$ and $sn(G)$ is at most c for any graph G . It is straightforward to show that $csn(G) \leq \Delta sn(G)$ for any graph G with maximum degree Δ . Moreover, it has been shown that there are graphs G for which $csn(G) = sn(G) + 1$. In [30], we investigate the question of the cost of the connectivity. We first provide new non-trivial upper and lower bounds for the cost of connectivity in the surveillance game. More precisely, we present a family of graphs G such that $csn(G) > sn(G) + 1$. Moreover, we prove that $csn(G) \leq \sqrt{sn(G)n}$ for any n -node graph G . While the gap between these bounds remains huge, it seems difficult to reduce it. We then define the online surveillance game where the observer has no a priori knowledge of the graph topology and discovers it little-by-little. This variant, which fits better the prefetching motivation, is a restriction of the connected variant. Unfortunately, we show that no algorithm for solving the online surveillance game has competitive ratio better than $\Omega(\Delta)$. That is, while interesting, this variant does not help to obtain better upper bounds for the connected variant. We finally answer an open question [68] by proving that deciding if the surveillance number of a digraph with maximum degree 6 is at most 2 is NP-hard.

7.2.3. Distributed Algorithms

7.2.3.1. Allowing each Node to Communicate only once in a Distributed System: Shared Whiteboard Models

In [21] we study distributed algorithms on massive graphs where links represent a particular relationship between nodes (for instance, nodes may represent phone numbers and links may indicate telephone calls). Since such graphs are massive they need to be processed in a distributed way. When computing graph-theoretic properties, nodes become natural units for distributed computation. Links do not necessarily represent communication channels between the computing units and therefore do not restrict the communication flow. Our goal is to model and analyze the computational power of such distributed systems where one computing unit is assigned to each node. Communication takes place on a whiteboard where each node is allowed to write at most one message. Every node can read the contents of the whiteboard and, when activated, can write one small message based on its local knowledge. When the protocol terminates its output is computed from the final contents of the whiteboard. We describe four synchronization models for accessing the whiteboard. We show that message size and synchronization power constitute two orthogonal hierarchies for these systems. We exhibit problems that separate these models, i.e., that can be solved in one model but not in a weaker one, even with increased message size. These problems are related to maximal independent set and connectivity. We also exhibit problems that require a given message size independently of the synchronization model.

7.2.3.2. Computing on Rings by Oblivious Robots: a Unified Approach for Different Tasks

A set of autonomous robots have to collaborate in order to accomplish a common task in a ring-topology where neither nodes nor edges are labeled (that is, the ring is anonymous). In [36], we present a unified approach to solve three important problems: the exclusive perpetual exploration, the exclusive perpetual clearing, and the gathering problems. In the first problem, each robot aims at visiting each node infinitely often while avoiding that two robots occupy a same node (exclusivity property); in exclusive perpetual clearing (also known as searching), the team of robots aims at clearing the whole ring infinitely often (an edge is cleared if it is traversed by a robot or if both its endpoints are occupied); and in the gathering problem, all robots must eventually occupy the same node. We investigate these tasks in the Look-Compute-Move model where the robots cannot communicate but can perceive the positions of other robots. Each robot is equipped with visibility sensors and motion actuators, and it operates in asynchronous cycles. In each cycle, a robot takes a snapshot of the current global configuration (Look), then, based on the perceived configuration, takes a decision to stay idle or to move to one of its adjacent nodes (Compute), and in the latter case it eventually moves to this neighbor (Move). Moreover, robots are endowed with very weak capabilities. Namely, they are anonymous, asynchronous, oblivious, uniform (execute the same algorithm) and have no common sense of

orientation. In this setting, we devise algorithms that, starting from an exclusive and rigid (i.e. aperiodic and asymmetric) configuration, solve the three above problems in anonymous ring-topologies.

7.2.4. Miscellaneous

7.2.4.1. Finding Paths in Grids with Forbidden Transitions

A transition in a graph is a pair of adjacent edges. Given a graph $G = (V, E)$, a set of forbidden transitions $F \subseteq E \times E$ and two vertices $s, t \in V$, we study in [64], [45], [46], [14] the problem of finding a path from s to t which uses none of the forbidden transitions of F . This means that it is forbidden for the path to consecutively use two edges forming a pair in F . The study of this problem is motivated by routing in road networks in which forbidden transitions are associated to prohibited turns as well as routing in optical networks with asymmetric nodes, which are nodes where a signal on an ingress port can only reach a subset of egress ports. If the path is not required to be elementary, the problem can be solved in polynomial time. On the other side, if the path has to be elementary, the problem is known to be NP-complete in general graphs [69]. In [45], we study the problem of finding an elementary path avoiding forbidden transitions in planar graphs. We prove that the problem is NP-complete in planar graphs and particularly in grids. In addition, we show that the problem can be solved in polynomial time in graphs with bounded treewidth. More precisely, we show that there is an algorithm which solves the problem in time $O((3\Delta(k+1))2k+4n)$ in n -node graphs with treewidth at most k and maximum degree Δ .

7.3. Graph theory

Participants: Nathann Cohen, Frédéric Havet.

7.3.1. Graph Colouring

7.3.1.1. Steinberg-like Theorems for Backbone Colouring

Motivated by some channel assignment problem, we study the following variation of graph colouring problem. A function $f : V(G) \rightarrow \{1, \dots, k\}$ is a (proper) k -colouring of G if $|f(u) - f(v)| \geq 1$, for every edge $uv \in E(G)$. The chromatic number $\chi(G)$ is the smallest integer k for which there exists a proper k -colouring of G . Given a graph G and a subgraph H of G , a circular q -backbone k -colouring c of (G, H) is a k -colouring of G such that $q \leq |c(u) - c(v)| \leq k - q$, for each edge $uv \in E(H)$. The circular q -backbone chromatic number of a graph pair (G, H) , denoted $CBC_q(G, H)$, is the minimum k such that (G, H) admits a circular q -backbone k -colouring. In [19], we first show that if G is a planar graph containing no cycle on 4 or 5 vertices and $H \subseteq G$ is a forest, then $CBC_2(G, H) \leq 7$. Then, we prove that if $H \subseteq G$ is a forest whose connected components are paths, then $CBC_2(G, H) \leq 6$.

7.3.1.2. Complexity of Greedy Edge-colouring

The Grundy index of a graph $G = (V, E)$ is the greatest number of colours that the greedy edge-colouring algorithm can use on G . In [33], we prove that the problem of determining the Grundy index of a graph $G = (V, E)$ is NP-hard for general graphs. We also show that this problem is polynomial-time solvable for caterpillars. More specifically, we prove that the Grundy index of a caterpillar is $\Delta(G)$ or $\Delta(G) + 1$ and present a polynomial-time algorithm to determine it exactly.

7.3.1.3. Proper Orientation Number

An *orientation* of a graph G is a digraph D obtained from G by replacing each edge by exactly one of the two possible arcs with the same endvertices. For each $v \in V(G)$, the *indegree* of v in D , denoted by $d_D^-(v)$, is the number of arcs with head v in D . An orientation D of G is *proper* if $d_D^-(u) \neq d_D^-(v)$, for all $uv \in E(G)$. The *proper orientation number* of a graph G , denoted by $\vec{\chi}(G)$, is the minimum of the maximum indegree over all its proper orientations. It is well-known that $\chi(G) \leq \vec{\chi}(G) + 1 \leq \Delta(G) + 1$, for every graph G , where $\chi(G)$ and $\Delta(G)$ denotes the chromatic number and the maximum degree of G . In other words, the proper orientation number (plus one) is an upper bound on the chromatic number which is tighter than the maximum degree.

In [17], we ask whether the proper orientation number is really a more accurate bound than the maximum degree in the following sense : does there exists a positive ϵ and such that $\vec{\chi}(G) \leq \epsilon \cdot \chi(G) + (1 - \epsilon)\Delta(G)$.

As an evidence to this, we prove that if G is bipartite (i.e. $\chi(G) \leq 2$) then $\vec{\chi}(G) \leq (\Delta(G) + \sqrt{\Delta(G)})/2 + 1$.

However, the proper orientation number has the drawback to be difficult to compute. We prove in [17] that deciding whether $\vec{\chi}(G) \leq \Delta(G) - 1$ is already an NP-complete problem on graphs with $\Delta(G) = k$, for every $k \geq 3$. We also show that it is NP-complete to decide whether $\vec{\chi}(G) \leq 2$, for planar *subcubic* graphs G . Moreover, we prove that it is NP-complete to decide whether $\vec{\chi}(G) \leq 3$, for planar bipartite graphs G with maximum degree 5.

Nevertheless, it might be interesting to bound the proper orientation number on some graph families. In particular, if we prove that for a graph with treewidth at most t , the proper orientation number is bounded by a function of t , this would imply that finding the proper orientation number of a graph with bounded treewidth is polynomial-time solvable. In [17] we prove $\vec{\chi}(G) \leq 4$ if G is a tree (or equivalently a graph with treewidth at most 1). In [53], we study the cacti which is a special class of graphs with treewidth at most 2. We prove that $\vec{\chi}(G) \leq 7$ for every cactus. We also prove that the bound 7 is tight by showing a cactus having no proper orientation with maximum indegree less than 7. We also prove that any planar claw-free graph has a proper orientation with maximum indegree at most 6 and that this bound can also be attained.

7.3.2. Subdivisions of Digraphs

An important result in the Roberston and Seymour minor theory is the polynomial-time algorithm to solve the so-called Linkage Problem. This implies in particular, that for any fixed graph H , deciding whether a graph G contains a subdivision of H as a subgraph can be solved in polynomial time.

We consider the directed analogue F -subdivision problem, which is an analogue for directed graphs (i.e. digraphs). Given a directed graph D , does it contain a subdivision of a prescribed digraph F ? In [20], we give a number of examples of polynomial instances, several NP-completeness proofs as well as a number of conjectures and open problems. In [62], we give further support to several open conjectures and speculations about algorithmic complexity of finding F -subdivisions. In particular, up to 5 exceptions, we completely classify for which 4-vertex digraphs F , the F -subdivision problem is polynomial-time solvable and for which it is NP-complete. While all NP-hardness proofs are made by reduction from some version of the 2-linkage problem in digraphs, some of the polynomial-time solvable cases involve relatively complicated algorithms.

7.4. Applications to Other Domains

Participants: Christelle Caillouet, David Coudert, Nicolas Nisse.

7.4.1. Unveiling Contacts within Macro-molecular assemblies by solving Minimum Weight Connectivity Inference Problems

Consider a set of oligomers listing the subunits involved in sub-complexes of a macro-molecular assembly, obtained e.g. using native mass spectrometry or affinity purification. Given these oligomers, connectivity inference (CI) consists in finding the most plausible contacts between these subunits, and minimum connectivity inference (MCI) is the variant consisting in finding a set of contacts of smallest cardinality. MCI problems avoid speculating on the total number of contacts, but yield a subset of all contacts and do not allow exploiting a priori information on the likelihood of individual contacts. In this context, we present in [15] two novel algorithms, MILP-W and MILP-WB. The former solves the minimum weight connectivity inference (MWCI), an optimization problem whose criterion mixes the number of contacts and their likelihood. The latter uses the former in a bootstrap fashion, to improve the sensitivity and the specificity of solution sets. Experiments on three systems (yeast exosome, yeast proteasome lid, human eIF3), for which reference contacts are known (crystal structure, cryo electron microscopy, cross-linking), show that our algorithms predict contacts with high specificity and sensitivity, yielding a very significant improvement over previous work, typically a twofold increase in sensitivity. The software accompanying this paper is made available, and should prove of ubiquitous interest whenever connectivity inference from oligomers is faced.

7.4.2. Recovery of Disrupted Airline Operations using k -Maximum Matching in Graphs

When an aircraft is approaching an airport, it gets a short time interval (called *slot*) that it can use to land. If the landing of the aircraft is delayed (because of bad weather, or if it arrives late, or if other aircrafts have to land first), it loses its slot and Air traffic controllers have to assign it a new slot. However, slots for landing are a scarce resource of the airports and, to avoid that an aircraft waits too much time, Air traffic controllers have to regularly modify the assignment of the slots of the aircrafts. Unfortunately, for legal and economical reasons, Air traffic controllers can modify the slot-assignment only using two kind of operations: either assign to aircraft A a slot that was free, or give to A the slot of another aircraft B and assign to B a free slot. The problem is then the following. Let $k \geq 1$ be an odd integer and let G be a graph and M be a matching (set of pairwise disjoint edges) of G . What is the maximum size of a matching that can be obtained from M by using only augmenting paths of length at most k ? Moreover, how to compute such a maximum matching? This problem has already been studied in the context of wireless networks, mainly because it provides a simple approximation for the classical matching problem. We prove in [65], [49] that this problem can be solved in polynomial-time when $k \leq 3$. Then, we show that, for any odd integer $k \geq 5$, the problem is NP-complete in planar bipartite graphs with maximum degree at most 3.

7.4.3. Inference of Curvilinear Structure based on Learning a Ranking Function and Graph Theory

To detect curvilinear structures in natural images, we propose in [63] a novel ranking learning system and an abstract curvilinear shape inference algorithm based on graph theory. We analyze the curvilinear structures as a set of small line segments. In this work, the rankings of the line segments are exploited to systematize the topological feature of the curvilinear structures. Structured Support Vector Machine is employed to learn the ranking function that predicts the correspondence of the given line segments and the latent curvilinear structures. We first extract curvilinear features using morphological profiles and steerable filtering responses. Also, we propose an orientation-aware feature descriptor and a feature grouping operator to improve the structural integrity during the learning process. To infer the curvilinear structure, we build a graph based on the output rankings of the line segments. We progressively reconstruct the curvilinear structure by looking for paths between remote vertices in the graph. Experimental results show that the proposed algorithm faithfully detects the curvilinear structures within various datasets.

7.4.4. Web Transparency for Complex Targeting: Algorithms, Limits, and Tradeoffs

Big Data promises important societal progress but exacerbates the need for due process and accountability. Companies and institutions can now discriminate between users at an individual level using collected data or past behavior. Worse, today they can do so in near perfect opacity. The nascent field of web transparency aims to develop the tools and methods necessary to reveal how information is used, however today it lacks robust tools that let users and investigators identify targeting using multiple inputs. In [67], [38], we formalize for the first time the problem of detecting and identifying targeting on combinations of inputs and provide the first algorithm that is asymptotically exact. This algorithm is designed to serve as a theoretical foundational block to build future scalable and robust web transparency tools. It offers three key properties. First, our algorithm is service agnostic and applies to a variety of settings under a broad set of assumptions. Second, our algorithm's analysis delineates a theoretical detection limit that characterizes which forms of targeting can be distinguished from noise and which cannot. Third, our algorithm establishes fundamental tradeoffs that lead the way to new metrics for the science of web transparency. Understanding the tradeoff between effective targeting and targeting concealment lets us determine under which conditions predatory targeting can be made unprofitable by transparency tools.

CTRL-A Team

7. New Results

7.1. Discrete control and reactive language support

Participants: Gwenaël Delaval, Eric Rutten, Stéphane Mocanu, Alia Hajjar, Abdoul-Razak Hassimi Harouna.

Concerning language support, we have designed and implemented BZR, a mixed imperative/declarative programming language: declarative contracts are enforced upon imperatively described behaviors (see 6.1). The semantics of the language uses the notion of Discrete Controller Synthesis (DCS) [5]. This work is done in close cooperation with the Inria team Sumo at Inria Rennes (H. Marchand).

New results concern the master internship of Alia Hajjar, co-directed by Gwenaël Delaval and Stéphane Mocanu, on the subject of Application of control of reactive environments and probabilistic models on Transactional Memory. Multiprocessor environments which use concurrent programs and data structures showed the need of techniques to organize the usage of the shared structures, to reduce the unpredicted delay and reduce the contention between concurrent processors. Transactional Memory (TM) is a programming model that eases development of concurrent applications. Concurrent programming causes conflicts and TM is a way to resolve these conflicts with the transaction paradigm. To control conflict, techniques are provided to optimize (identify the best) degree of parallelism. In this framework, the aim is to control the TM system by adapting the degree of parallelism in order to maximize the throughput, i.e., number of committed transactions per time. The main objective is to minimize the execution time of a parallel application, thus maximize the throughput. During this master's thesis, the behavior of a multithreaded TM environment has been modeled as a stochastic discrete event system. The Heptagon/BZR language has then been used to implement this model for simulation, and evaluation of control strategies.

Ongoing work concerns aspects of compilation and debugging and exploring the notion of adaptive discrete control, which is yet an open question in discrete control in contrast to the well-known adaptive continuous control.

Another activity related to discrete control is or work with Leiden University and CWI (N. Khakpour, now at Linnaeus U., and F. Arbab) on enforcing correctness of the behavior of an adaptive software system during dynamic adaptation is an important challenge along the way to realize correct adaptive systems [11].

7.2. Design and programming

7.2.1. Component-based approaches

Participants: Frederico Alvares de Oliveira Junior, Eric Rutten.

Architecting in the context of variability has become a real need in today's software development. Modern software systems and their architecture must adapt dynamically to events coming from the environment (e.g., workload requested by users, changes in functionality) and the execution platform (e.g., resource availability). Component-based architectures have shown to be very suited for self-adaptation especially with their dynamical reconfiguration capabilities. However, existing solutions for reconfiguration often rely on low level, imperative, and non formal languages. We have defined Ctrl-F, a domain-specific language whose objective is to provide high-level support for describing adaptation behaviors and policies in component-based architectures. It relies on reactive programming for formal verification and control of reconfigurations. We integrate Ctrl-F with the FraSCAti Service Component Architecture middleware platform, and apply it to the Znn.com self-adaptive case study [20], [15], [14], [18].

We work on the topic in cooperation with the Spirals Inria team at Inria Lille (L. Seinturier). It constitutes a follow-up on previous work in the ANR Minalogic project MIND, industrializing the Fractal component-based framework, with a continuation of contacts with ST Microelectronics (V. Bertin). Our integration of BZR and Fractal [4], [2] is at the basis of our current work.

7.2.2. Rule-based systems

Participants: Adja Sylla, Eric Rutten.

We are starting a cooperation with CEA LETI/DACLE on the topic of a high-level language for safe rule-based programming in the LINC platform. The general context is that of the runtime redeployment of distributed applications, for example managing smart buildings. Motivations for redeployment can be diverse: load balancing, energy saving, upgrading, or fault tolerance. Redeployment involves changing the set of components in presence, or migrating them. The basic functionalities enabling to start, stop, migrate, or clone components, and the control managing their safe coordination, will have to be designed in the LINC middleware developed at CEA.

The transactional nature of the LINC platform insures the correct execution of each of the rules constituting the program, but there still is a need to insure the safety of their coordination, and of the behavior resulting from their sequential execution. For example, in the smart environments application domain, we must insure safety of control decisions, so that all the configurations that can be reached are safe, as well as the sequences of actions in switching between them. For this we will rely on automata-based models and control, using the BZR language, and integrating it in a domains specific language. Our work builds upon preliminary results involving colored Petri nets models [17].

The PhD of Adja Sylla at CEA on this topic is co-advised with F. Pacull and M. Louvel.

7.3. Infrastructure-level support

We apply the results of the previous axes of the team's activity to a range of infrastructures of different natures, but sharing a transversal problem of reconfiguration control design. From this very diversity of validations and experiences, we draw a synthesis of the whole approach [13], towards a general view of Feedback Control as MAPE-K loop in Autonomic Computing [21].

7.3.1. Autonomic Cloud and Big-Data systems

7.3.1.1. Coordination in multiple-loop autonomic Cloud systems

Participants: Soguy Gueye, Gwenaël Delaval, Eric Rutten.

Complex computing systems are increasingly self-adaptive, with an autonomic computing approach for their administration. Real systems require the co-existence of multiple autonomic management loops, each complex to design. However their uncoordinated co-existence leads to performance degradation and possibly to inconsistency. There is a need for methodological supports facilitating the coordination of multiple autonomic managers. To tackle this problem, we take a global view and underscore that Autonomic Management Systems (AMS) are intrinsically reactive, as they react to flows of monitoring data by emitting flows of reconfiguration actions. Therefore we propose a new approach for the design of AMSs, based on synchronous programming and discrete controller synthesis techniques. They provide us with high-level languages for modeling the system to manage, as well as means for statically guaranteeing the absence of logical coordination problems. Hence, they suit our main contribution, which is to obtain guarantees at design time about the absence of logical inconsistencies in the taken decisions. We detail our approach, illustrate it by designing an AMS for a realistic multi-tier application, and evaluate its practicality with an implementation [10].

In order to coordinate managers without breaking their natural modularity, we address the problem with a method stressing modularity, and focusing on the discrete control of the interactions of managers. We make proposals for the distributed execution of modular controllers, first in synchronized way, and then relaxing this synchronization. We apply and validate our method on a multi-loop multi-tier system in a data-center [16].

We addressed these problems in the context of the ANR project Ctrl-Green, in cooperation with LIG (N. de Palma) in the framework of the PhD of S. Gueye and the post-doc of N. Berthier.

7.3.1.2. Control for Big data

Participants: Bogdan Robu [Gipsa-lab], Mihaly Berekmeri [Gipsa-lab], Nicolas Marchand [Gipsa-lab].

Companies have a fast growing amounts of data to process and store, a data explosion is happening next to us. Currently one of the most common approaches to treat these vast data quantities is the MapReduce parallel programming paradigm. While it's use is widespread in the industry, ensuring performance constraints, while also minimizing costs, provides considerable challenges. To deal with these issues we propose a control theoretical approach, based on techniques that have already proved their usefulness in the control community. We developed an algorithm to create the first linear dynamic model for a Big Data MapReduce Cloud system, running a concurrent workload. Furthermore we identify two important control use cases: relaxed performance - minimal resource and strict performance. We developed the first feedback control mechanism for such systems. Then to minimize the number of control actuations, an event-based feedback controller was also introduced. Furthermore to address the strict performance challenges a feedforward controller that efficiently suppresses the effects of large workload size variations is developed. On top of this issues an optimal predictive control which deals with concurrent objectives (dependability and performance) is implemented. The approach is validated online in a benchmark running in a real 60 node MapReduce cluster, using a data intensive Business Intelligence [22], [23].

This work is performed in cooperation with LIG (S. Bouchenak) in the framework of the PhD of M. Berekmeri.

7.3.2. *Reconfiguration control in DPR FPGA*

Participant: Eric Rutten.

Dynamically reconfigurable hardware has been identified as a promising solution for the design of energy efficient embedded systems. However, its adoption is limited by the costly design effort including verification and validation, which is even more complex than for non dynamically reconfigurable systems. We worked on this topic in the context of a ensign environment, developed in the framework of the ANR project Famous, in cooperation with LabSticc in Lorient and Inria Lille (DaRT team) [12]. We proposed a tool-supported formal method to automatically design a correct-by-construction control of the reconfiguration. By representing system behaviors with automata, we exploit automated algorithms to synthesize controllers that safely enforce reconfiguration strategies formulated as properties to be satisfied by control. We design generic modeling patterns for a class of reconfigurable architectures, taking into account both hardware architecture and applications, as well as relevant control objectives. We validate our approach on two case studies implemented on FPGAs [1].

We are currently valorizing results in more publications [12], [9], and extending the use of control techniques by evaluating the new tool ReaX developed at Inria Rennes (Sumo).

We are starting a new ANR project called HPeC, within which some of these topics will be extended, especially regarding hierarchical and modular control, and logico-numeric aspects.

7.3.3. *Autonomic memory management in HPC*

Participants: Naweiluo Zhou, Gwenaël Delaval, Bogdan Robu, Eric Rutten.

Parallel programs need to manage the time trade-off between synchronization and computation. A high parallelism may decrease computing time but meanwhile increase synchronization cost among threads. Software Transactional Memory (STM) has emerged as a promising technique, which bypasses locks, to address synchronization issues through transactions. A way to reduce conflicts is by adjusting the parallelism, as a suitable parallelism can maximize program performance. However, there is no universal rule to decide the best parallelism for a program from an offline view. Furthermore, an offline tuning is costly and error-prone. Hence, it becomes necessary to adopt a dynamical tuning-configuration strategy to better manage a STM system. Autonomic control techniques begin to receive attention in computing systems recently. Control technologies offer designers a framework of methods and techniques to build autonomic systems with well-mastered behaviors. The key idea of autonomic control is to implement feedback control loops to design safe, efficient and predictable controllers, which enable monitoring and adjusting controlled systems dynamically while keeping overhead low. We propose to design feedback control loops to automate the choice of parallelism at runtime and diminish program execution time.

In the context of the action-team HPES of the Labex Persyval-lab⁰ (see 9.1), this work is performed in cooperation with LIG (J.F. Méhaut) in the framework of the PhD of N. Zhou.

7.3.4. Control of smart environments

Participants: Adja Sylla, Mengxuan Zhao, Eric Rutten, Hassane Alla [Gipsa-lab].

7.3.4.1. Generic supervision architecture

New application domains of control, such as in the Internet of Things (IoT) and Smart Environments, require generic control rules enabling the systematization and the automation of the controller synthesis. We worked on an approach for the generation of Discrete Supervisory Controllers for these applications. A general modeling framework is proposed for the application domain of smart home. We formalize the design of the environment manager as a Discrete Controller Synthesis (DCS) problem, w.r.t. multiple constraints and objectives, for example logical issues of mutual exclusion, bounding of power peaks. We validate our models and manager computations with the BZR language and an experimental simulator This work was performed in cooperation with Orange labs (G. Privat) in the framework of the Cifre PhD of M. Zhao [8].

7.3.4.2. Rule-based specification

In the context of IoT applications like smart home environments, the rules for programming in the LINC framework are used as a flexible tool to govern the relations between sensors and actuators. Runtime coordination and formal analysis becomes a necessity to avoid side effects mainly when applications are critical. In cooperation with CEA LETI/DACLE, we are working on a case study for safe applications development in IoT and smart home environments [17].

⁰<https://persyval-lab.org/en/sites/hpes>

DANTE Project-Team

7. New Results

7.1. Graph & Signal Processing

Participants: Paulo Gonçalves Andrade, Éric Fleury, Benjamin Girault, Sarra Ben Alaya.

Isometric Graph shift operator. In [14], [40], we proposed a new shift operator for graph signals, enforcing that our operator is isometric. Doing so, we ensure that as many properties of the time shift as possible get carried over. Finally, we show that our operator behaves reasonably for graph signals.

Stationary graph signals. We extended the concept of stationary temporal signals to stationary graph signals [24]. We introduced the concept of strict sense stationary and wide sense stationary graph signals as a statistical invariance through an isometric graph translation. Using these definitions, we proposed a spectral characterisation of WSS graph signals allowing to study stationarity using only the spectral components of a graph signal. Finally, we applied this characterisation to a synthetic graph in order to study a few important stochastic graph signals. Also, using geographic data, we analysed data from a graph set of weather stations and showed evidence of stationarity in the temperature signal [36].

Community mining with graph filters for correlation matrices. Communities are an important type of structure in networks. Graph filters, such as wavelet filter-banks, have been used to detect such communities as groups of nodes more densely connected together than with the outsiders. When dealing with times series, it is possible to build a relational network based on the correlation matrix. However, in such a network, weights assigned to each edge have different properties than those of usual adjacency matrices. As a result, classical community detection methods based on modularity optimisation are not consistent and the modularity needs to be redefined to take into account the structure of the correlation from random matrix theory. In our contribution [34], we addressed how to detect communities from correlation matrices, by filtering global modes and random parts using properties that are specific to the distribution of correlation eigenvalues. Based on a Louvain approach, an algorithm to detect multiscale communities was also developed, which yields a weighted hierarchy of communities. The implementation of the method using graph filters was also discussed.

A strong Tauberian theorem for characteristic functions. In [20], we showed that a characteristic function which can be approximated at 0 by any polynomial of order n is actually n -times differentiable at 0. This fact is exploited to strengthen a tauberian-type result by Lukacs and provides the theoretical basis for a wavelet based non-parametric estimator of the tail index of a distribution. This work is a technical improvement of our previous contribution [53].

Fractal Analysis of Fetal Heart Rate Variability. The fetal heart rate (FHR) is commonly monitored during labor to detect early fetal acidosis. FHR variability is traditionally investigated using Fourier transform, often with adult predefined frequency band powers and the corresponding LF/HF ratio. However, fetal conditions differ from adults and modify spectrum repartition along frequencies. The study we reported in [12] questioned the arbitrariness definition and relevance of the frequency band splitting procedure, and thus of the calculation of the underlying LF/HF ratio, as efficient tools for characterising intrapartum FHR variability. Then, we showed that the intrapartum FHR is characterised by fractal temporal dynamics and promotes the Hurst parameter as a potential marker of fetal acidosis. This parameter preserves the intuition of a power frequency balance, while avoiding the frequency band splitting procedure and thus the arbitrary choice of a frequency separating bands. The study also shows that extending the frequency range covered by the adult-based bands to higher and lower frequencies permits the Hurst parameter to achieve better performance for identifying fetal acidosis.

7.2. Performance analysis and networks protocols

Participants: Paulo Gonçalves Andrade, Thomas Begin, Anthony Busson, Isabelle Guérin Lassous, Laurent Reynaud, Thiago Wanderley Matos de Abreu.

Global computing-network-visualisation. The PetaFlow application aims to contribute to the use of high performance computational resources for the benefit of society. To this goal the emergence of adequate information and communication technologies with respect to high performance computing-networking-visualisation and their mutual awareness is required. In the work published in [5], we present the developed technology and the algorithms that we applied to a real global peta-scale data intensive scientific problem with social and medical importance, i.e. human upper airflow modeling.

Performance analysis of multi-hop flows in IEEE 802.11 networks Multi-hop wireless networks are often regarded as a promising means to extend the limited coverage area offered by WLANs. However, they are usually associated with poor and uncertain performance in terms of available bandwidth and packet losses, which clearly stands as a limitation to their use. In [7], we consider the performance evaluation of a multi-hop path (also called chain), based on the IEEE 802.11 DCF. The proposed modeling framework is constructive and versatile, so that it can handle various types of multi-hop wireless paths, including scenarios with two flows in opposite directions, and topologies where nodes are exposed to the well-known hidden node problem. The models derived from our framework are conceptually simple, easy to implement and produce generally accurate results for the attained goodput of flows, as well as the datagram loss probability. Typical relative errors for these two quantities are below a few percent. Also, fundamental phenomena occurring in multi-hop wireless networks such as performance collapse and starvation, are well captured by the models.

Passive Measurement-based Estimator for the Standard Deviation of the End-to-End Delay.

Emerging architectures for computer networks such as SDN aim at offering a better handling of flows with stringent requirements of QoS. On the one hand, operators would benefit from a detailed description of common network performance (e.g., end-to-end delay and end-to-end loss ratio) including their first two moments, namely mean and standard deviation. Indeed, for many applications, the variability in the end-to-end delay (e.g., jitter) deeply affects the actual QoS experienced by a flow. On the other hand, the cost and nuisance associated with the instrumentation, the measurements, and the computations must be kept as low as possible. This typically prevents the availability of end-to-end measurements. In [30], we propose an algorithm to estimate the second moment of the end-to-end delay experienced by the packets of a flow based only on delay measurements locally collected by the network nodes. Our solution estimates the standard deviation of the end-to-end delay in an easy and computationally efficient way. Based on thousands of simulations using a real-life trace, our solution is found to be accurate, typically differing by only a few percent from the actual value of the standard deviation of the end-to-end delay.

Design of a force-based controlled mobility on aerial vehicles for pest management. *Vespa velutina*, also known as the Asian hornet, is considered as an invasive species out of its native zone. In particular, since it preys on honey bees, its recent progression in Europe could soon pose a significant risk to the local apiculture activity. European beekeepers are therefore investigating adapted control strategies, including *V. velutina* nest destruction. Unfortunately, nest location pinpointing generally follows a manual process which can prove tedious, time-consuming and inaccurate. In [31], we propose the use of a network of micro aerial vehicles featuring autonomous and cooperative flight capabilities. We describe an adapted controlled mobility strategy and detail the design of our Virtual Force Protocol (VFP) which allows a swarm of vehicles to track and follow hornets to their nests, while maintaining connectivity through a wireless multi-hop communication route with a remote ground station used to store applicative data such as hornet trajectory and vehicle telemetry. In order to achieve the mission objectives with a minimum of vehicles, we identify through simulations appropriate value for the key parameters of VFP and discuss the obtained network performance.

Channel assignment in IEEE 802.11-based substitution networks. A substitution network is a rapidly deployable wireless network that provides a backup solution to quickly react to failures

on an existing network. We assume that the substitution network uses Wi-Fi technology and that wireless routers are equipped with several Wi-Fi cards. The problem, addressed in this work, deals with the channel assignment to these wireless interfaces. In this particular context, there is only one source-destination pair for which paths are known in advance. It is then possible to derive an objective function, function of the channel assignment, that very precisely reflects the overall throughput that can be achieved in this network. This problem is formulated through a linear optimization problem for which we propose different heuristics. Simulation results, performed with ns-3, consider several scenarios, and compare our heuristics to the optimum. Simulations show that, with only a few wireless cards, the throughput is significantly increased. Also, we show that the objective function fits to the throughput measured with ns-3.

Performance evaluation and message dissemination in vehicular networks. Vehicular Ad-Hoc Network (VANET) is becoming a promising technology for improving the efficiency and the safety of Intelligent Transportation Systems (ITS). Smart vehicles are expected to continuously exchange a huge amount of data either through safety or non-safety messages dedicated for road safety or infotainment and passenger comfort applications, respectively. In this context we proposed two contributions: the estimation of the capacity offered by the wireless network [13] in order to dimension the applications, and the proposal of an efficient message dissemination protocol [25].

Performance Evaluation of Cloud Computing Centers with General Arrivals and Service. Cloud providers need to size their systems to determine the right amount of resources to allocate as a function of customer's needs so as to meet their SLAs (Service Level Agreement), while at the same time minimizing their costs and energy use. Queueing theory based tools are a natural choice when dealing with performance aspects of the QoS (Quality of Service) part of the SLA and forecasting resource utilization. The characteristics of a cloud center lead to a queueing system with multiple servers (nodes) in which there is potentially a very large number of servers and both the arrival and service process can exhibit high variability. We propose to use a G/G/c-like model to represent a cloud system and assess expected performance indices. Given the potentially high number of servers in a cloud system, we present an efficient, fast and easy-to-implement approximate solution. We have extensively validated our approximation against discrete-event simulation for several QoS performance metrics such as task response time and blocking probability with excellent results. We apply our approach to examples of system sizing and our examples clearly demonstrate the importance of taking into account the variability of the tasks arrivals and thus expose the risk of under- or over-provisioning if one relies on a model with Poisson assumptions [8].

Prediction of the System Performance from components models. In this paper we consider the problem of combining calibrated performance models of system components in order to predict overall system performance. We focus on open workload system models, in which, under certain conditions, obtaining and validating the overall system performance measures can be a simple application of Little's law. We discuss the conditions of applicability of such a simple validation methodology, including examples of successful application, as well as examples where this approach fails. Additionally, we propose to analyze the deviations between the model predictions and system measurements, so as to decide if they correspond to "measurement noise" or if an important system component has not been correctly represented. This approach can be used as an aid in the design of validated system performance models [26].

7.3. Modeling of Dynamics of Complex Networks

Participants: Christophe Crespelle, Éric Fleury, Márton Karsai, Yannick Leo, Matteo Morini.

Non-Altering Time Scales for Aggregation of Dynamic Networks into Series of Graphs [29] Many dynamic networks coming from real-world contexts are *link streams*, i.e. a finite collection of triplets (u, v, t) where u and v are two nodes having a link between them at time t . A great number of studies on these objects start by aggregating the data on disjoint time windows of length Δ in order to obtain a series of graphs on which are made all subsequent analyses. Here we are concerned

with the impact of the chosen Δ on the obtained graph series. We address the fundamental question of knowing whether a series of graphs formed using a given Δ faithfully describes the original link stream. We answer the question by showing that such dynamic networks exhibit a threshold for Δ , which we call the *saturation scale*, beyond which the properties of propagation of the link stream are altered, while they are mostly preserved before. We design an automatic method to determine the saturation scale of any link stream, which we apply and validate on several real-world datasets.

Termination of the Iterated Strong-Factor Operator on Multipartite Graphs [10] The clean-factor operator is a multipartite graph operator that has been introduced in the context of complex network modelling. Here, we consider a less constrained variation of the clean-factor operator, named strong-factor operator, and we prove that, as for the clean-factor operator, the iteration of the strong-factor operator always terminates, independently of the graph given as input. Obtaining termination for all graphs using minimal constraints on the definition of the operator is crucial for the modelling purposes for which the clean-factor operator has been introduced. Moreover we show that the relaxation of constraints we operate not only preserves termination but also preserves the termination time, in the sense that the strong-factor series always terminates before the clean-factor series.

On the Termination of Some Biclique Operators on Multipartite Graphs [9] We define a new graph operator, called the *weak-factor graph*, which comes from the context of complex network modelling. The weak-factor operator is close to the well-known clique-graph operator but it rather operates in terms of bicliques in a multipartite graph. We address the problem of the termination of the series of graphs obtained by iteratively applying the weak-factor operator starting from a given input graph. As for the clique-graph operator, it turns out that some graphs give rise to series that do not terminate. Therefore, we design a slight variation of the weak-factor operator, called *clean-factor*, and prove that its associated series terminates for all input graphs. In addition, we show that the multipartite graph on which the series terminates has a very nice combinatorial structure: we exhibit a bijection between its vertices and the chains of the inclusion order on the intersections of the maximal cliques of the input graph.

Directed Cartesian-Product Decomposition [11]. In this paper, we design an algorithm that, given a directed graph G and the Cartesian-product decomposition of its underlying undirected graph \tilde{G} , produces the directed Cartesian-product decomposition of G in linear time. This is the first time that the linear complexity is achieved for this problem, which has two major consequences. Firstly, it shows that the directed and undirected versions of the Cartesian-product decomposition of graphs are linear-time equivalent problems. And secondly, as there already exists a linear-time algorithm for solving the undirected version of the problem, combined together, it provides the first linear-time algorithm for computing the directed Cartesian-product decomposition of a directed graph.

An $O(n^2)$ time Algorithm for the Minimal Permutation Completion Problem [28] We provide an $O(n^2)$ time algorithm computing a minimal permutation completion of an arbitrary graph $G = (V, E)$, i.e., a permutation graph $H = (V, F)$ on the same vertex set, such that $E \subseteq F$ and F is inclusion-minimal among all possibilities.

Linearity is Strictly More Powerful than Contiguity for Encoding Graphs [27] Linearity and contiguity are two parameters devoted to graph encoding. Linearity is a generalisation of contiguity in the sense that every encoding achieving contiguity k induces an encoding achieving linearity k , both encoding having size $\Theta(k.n)$, where n is the number of vertices of G . In this paper, we prove that linearity is a strictly more powerful encoding than contiguity, i.e. there exists some graph family such that the linearity is asymptotically negligible in front of the contiguity. We prove this by answering an open question asking for the worst case linearity of a cograph on n vertices: we provide an $O(\log n / \log \log n)$ upper bound which matches the previously known lower bound.

Socioeconomic correlations in communication networks [37], [38] In this work we study the socioeconomic structure of a communication network by combining mobile communication records and bank credit informations of a large number of individuals living in Mexico. We provide empirical evidences about present economic unbalances suggesting not only the distribution of wealth but also

the distribution of debts to follow the Pareto principle. Further we study the internal and interconnected structure of socioeconomic groups. Through a weighted core analysis we signal assortative correlations between people regarding their economic capacities, and show the existence of "rich-clubs" indicating present social stratification in the social structure. This project is ongoing with final results expected in 2016.

Detecting global bridges in networks [15] The identification of nodes occupying important positions in a network structure is crucial for the understanding of the associated real-world system. Usually, betweenness centrality is used to evaluate a node capacity to connect different graph regions. However, we argue here that this measure is not adapted for that task, as it gives equal weight to "local" centers (i.e. nodes of high degree central to a single region) and to "global" bridges, which connect different communities. This distinction is important as the roles of such nodes are different in terms of the local and global organisation of the network structure. In this paper we propose a decomposition of betweenness centrality into two terms, one highlighting the local contributions and the other the global ones. We call the latter bridgeness centrality and show that it is capable to specifically spot out global bridges. In addition, we introduce an effective algorithmic implementation of this measure and demonstrate its capability to identify global bridges in air transportation and scientific collaboration networks.

Collective attention in the age of (mis)information [17] We study, on a sample of 2.3 million individuals, how Facebook users consumed different information at the edge of political discussion and news during the last Italian electoral competition. Pages are categorized, according to their topics and the communities of interests they pertain to, in a) alternative information sources (diffusing topics that are neglected by science and main stream media); b) online political activism; and c) main stream media. We show that attention patterns are similar despite the different qualitative nature of the information, meaning that unsubstantiated claims (mainly conspiracy theories) reverberate for as long as other information. Finally, we categorize users according to their interaction patterns among the different topics and measure how a sample of this social ecosystem (1279 users) responded to the injection of 2788 false information posts. Our analysis reveals that users which are prominently interacting with alternative information sources (i.e. more exposed to unsubstantiated claims) are more prone to interact with false claims.

The Scaling of Human Contacts in Reaction-Diffusion Processes [22] We present new empirical evidence, based on millions of interactions on Twitter, confirming that human contacts scale with population sizes. We integrate such observations into a reaction-diffusion metapopulation framework providing an analytical expression for the global invasion threshold of a contagion process. Remarkably, the scaling of human contacts is found to facilitate the spreading dynamics. Our results show that the scaling properties of human interactions can significantly affect dynamical processes mediated by human contacts such as the spread of diseases, and ideas.

From calls to communities: a model for time varying social networks [16] Social interactions vary in time and appear to be driven by intrinsic mechanisms, which in turn shape the emerging structure of the social network. Large-scale empirical observations of social interaction structure have become possible only recently, and modelling their dynamics is an actual challenge. Here we propose a temporal network model which builds on the framework of activity-driven time-varying networks with memory. The model also integrates key mechanisms that drive the formation of social ties - social reinforcement, focal closure and cyclic closure, which have been shown to give rise to community structure and the global connectedness of the network. We compare the proposed model with a real-world time-varying network of mobile phone communication and show that they share several characteristics from heterogeneous degrees and weights to rich community structure. Further, the strong and weak ties that emerge from the model follow similar weight-topology correlations as real-world social networks, including the role of weak ties.

Kinetics of Social Contagion [21] Diffusion of information, behavioural patterns or innovations follows diverse pathways depending on a number of conditions, including the structure of the underlying social network, the sensitivity to peer pressure and the influence of media. Here we study

analytically and by simulations a general model that incorporates threshold mechanism capturing sensitivity to peer pressure, the effect of 'immune' nodes who never adopt, and a perpetual flow of external information. While any constant, non-zero rate of dynamically-introduced innovators leads to global spreading, the kinetics by which the asymptotic state is approached show rich behaviour. In particular we find that, as a function of the density of immune nodes, there is a transition from fast to slow spreading governed by entirely different mechanisms. This transition happens below the percolation threshold of fragmentation of the network, and has its origin in the competition between cascading behaviour induced by innovators and blocking of adoption due to immune nodes. This change is accompanied by a percolation transition of the induced clusters.

DIANA Project-Team

6. New Results

6.1. Service Transparency

6.1.1. From Network-level Measurements to Expected QoE: the Skype Use Case

Participants: Thierry Spetebroot, Nicolas Aguilera, Damien Saucez and Chadi Barakat.

Modern Internet applications rely on rich multimedia contents making the quality of experience (QoE) of end users sensitive to network conditions. Several models were developed in the literature to express QoE as a function of measurements carried out on the traffic of the applications themselves. In this contribution, we propose a new methodology based on machine learning able to link expected QoE to network and device level measurements outside the applications' traffic. This direct linking to network and device level measurements is important for the prediction of QoE. We prove the feasibility of the approach in the context of Skype. In particular, we derive and validate a model to predict the Skype QoE as a function of easily measurable network performance metrics. One can see our methodology as a new way of performing measurements in the Internet, where instead of expressing the expected performance in terms of network and device level measurements that only specialists can understand, we express performance in clear terms related to expected quality of experience for different applications. More details on this approach and on our application ACQUA can be found in section 5.1, in the paper summarizing the results [16] and on the application web page <http://team.inria.fr/diana/acqua/>.

6.1.2. Towards a General Solution for Detecting Traffic Differentiation at the Internet Access

Participants: Ricardo Ravaioli and Chadi Barakat.

In recent years network neutrality has been widely debated from both technical and economic points of view. Various cases of traffic differentiation at the Internet access have been reported throughout the last decade, in particular aimed at bandwidth consuming traffic flows. In this contribution we present a novel application-agnostic method for the detection of traffic differentiation, through which we are able to correctly identify where a shaper is located with respect to the user and evaluate whether it affected delays, packet losses or both. The tool we propose, ChkDiff, replays the user's own traffic in order to target routers at the first few hops from the user. By comparing the resulting flow delays and losses to the same router against one other, and analyzing the behaviour on the immediate router topology spawning from the user end-point, ChkDiff manages to detect instances of traffic shaping. This contribution is published in [15] where we provide a detailed description of the design of the tool for the case of upstream traffic, the technical issues it overcomes and a validation in controlled scenarios. It is the result of collaboration with the SIGNET group at I3S in the context of a PhD thesis funded by the UCN@SOPHIA Labex.

6.1.3. A Diagnostic Tool for Content-Centric Networks

Participant: Thierry Turetletti

In collaboration with our colleagues at NICT, Japan, we have proposed the Contrace tool for Measuring and Tracing Content-Centric Networks (CCNs). CCNs are fundamental evolutionary technologies that promise to form the cornerstone of the future Internet. The information flow in these networks is based on named data requesting, in-network caching, and forwarding – which are unique and can be independent of IP routing. As a result, common IP-based network tools such as ping and traceroute can neither trace a forwarding path in CCNs nor feasibly evaluate CCN performance. We designed "contrace," a network tool for CCNs (particularly, CCNx implementation running on top of IP) that can be used to investigate 1) the Round-Trip Time (RTT) between content forwarder and consumer, 2) the states of in-network cache per name prefix, and 3) the forwarding path information per name prefix. We report a series of experiments conducted using contrace on a CCN topology created on a local testbed and the GEANT network topology emulated by the Mini-CCNx emulator.

The results confirm that contrace is not only a useful tool for monitoring and operating a network, but also a helpful analysis tool for enhancing the design of CCNs. Further, contrace can report the number of received interests per cache or per chunk on the forwarding routers. This enables us to estimate the content popularity and design more effective cache control mechanisms in experimental networks (see our publication in the IEEE Communication Magazine [9]).

6.1.4. An efficient packet extraction tool for large experimentation traces

Participants: Thierry Turletti and Walid Dabbous

Network packet tracing has been used for many different purposes during the last few decades, such as network software debugging, networking performance analysis, forensic investigation, and so on. Meanwhile, the size of packet traces becomes larger, as the speed of network rapidly increases. Thus, to handle huge amounts of traces, we need not only more hardware resources, but also efficient software tools. However, traditional tools are inefficient at dealing with such big packet traces. In this work, we propose pcapWT, an efficient packet extraction tool for large traces. PcapWT provides fast packet lookup by indexing an original trace using a Wavelet Tree structure. In addition, it supports multi-threading for avoiding synchronous I/O and blocking system calls used for file processing, and it is particularly efficient on machines with SSD disks. PcapWT shows remarkable performance enhancements in comparison with traditional tools such as tcpdump and most recent tools such as pcapIndex in terms of index data size and packet extraction time. Our benchmark using large and complex traces shows that pcapWT reduces the index data size down below 1% of the volume of the original traces. Moreover, packet extraction performance is 20% better than with pcapIndex. Furthermore, when a small amount of packets are retrieved, pcapWT is hundreds of times faster than tcpdump. This work has been done in collaboration with our colleagues at Universidad Diego Portales (UDP) and Universidad de Chile and has been published in the Computer Networks journal [10].

6.1.5. Social Clicks: What and Who Gets Read on Twitter?

Participants: Maksym Gabielkov and Arnaud Legout

Online news domains increasingly rely on social media to drive traffic to their website. Yet we know surprisingly little about how social media conversation mentioning an online article actually generates a click to it. Posting behaviors, in contrast, have been fully or partially available and scrutinized over the years. While this has led to multiple assumptions on the diffusion of information, each were designed or validated while ignoring this important step. We made a large scale, validated and reproducible study of social clicks, that is also the first data of its kind, gathering a month of web visits to online resources that are located in 5 leading news domains and that are mentioned in the third largest social media by web referral (Twitter). Our dataset amounts to 2.8 million posts, together responsible for 75 billion potential views on this social media, and 9.6 million actual clicks to 59,088 unique resources. We design a reproducible methodology, carefully corrected its biases, enabling data sharing, future collection and validation. As we prove, properties of clicks and social media Click-Through-Rates (CTR) impact multiple aspects of information diffusion, all previously unknown. Secondary resources, that are not promoted through headlines and are responsible for the long tail of content popularity, generate more clicks both in absolute and relative terms. Social media attention is actually long-lived, in contrast with temporal evolution estimated from posts or impressions. The actual influence of an intermediary or a resource is poorly predicted by their posting behavior, but we show how that prediction can be made more precise. The results are reported in an article under submission, no report available yet.

6.1.6. ReCon: Revealing and Controlling PII Leaks in Mobile Network Traffic

Participant: Arnaud Legout

It is well known that apps running on mobile devices extensively track and leak users' personally identifiable information (PII); however, these users have little visibility into PII leaked through the network traffic generated by their devices, and have poor control over how, when and where that traffic is sent and handled by third parties. In this work, we present the design, implementation, and evaluation of ReCon: a cross-platform system that reveals PII leaks and gives users control over them without requiring any special privileges or custom OSes. ReCon leverages machine learning to reveal potential PII leaks by inspecting network traffic, and provides a visualization tool to empower users with the ability to control these leaks via blocking or

substitution of PII. We evaluate ReCon's effectiveness with measurements from controlled experiments using leaks from the 100 most popular iOS, Android, and Windows Phone apps, and via an user study with 92 participants. In this study, that was approved by the Inria Ethical Board (COERELE), we show that ReCon is accurate, efficient, and identifies a wider range of PII than previous approaches. The results are reported in an article under submission, no report available yet.

6.2. Open Network Architecture

6.2.1. *Storage on Wheels: Offloading Popular Contents Through a Vehicular Cloud*

Participants: Luigi Vigneri and Chadi Barakat.

The increasing demand for mobile data is overloading the cellular infrastructure. Small cells and edge caching is being explored as an alternative, but installation and maintenance costs for sufficient coverage are significant. In this work, we perform a preliminary study of an alternative architecture based on two main ideas: (i) using vehicles as mobile caches that can be accessed by user devices; compared to small cells, vehicles are more widespread and require lower costs; (ii) combining the mobility of vehicles with delayed content access to increase the number of cache hits (and reduce the load on the infrastructure). Contrary to standard DTN-type approaches, in our system max delays are guaranteed to be kept to a few minutes (beyond this deadline, the content is fetched from the infrastructure). We first propose an analytical framework to compute the optimal number of content replicas that one should cache, in order to minimize the infrastructure load. We then investigate how to optimally refresh these caches to introduce new contents, as well as to react to the temporal variability in content popularity. Simulations suggest that our vehicular cloud considerably reduces the infrastructure load in urban settings, assuming modest penetration rates and tolerable content access delays. This work is currently under submission; it is the result of collaboration with the Mobile Communications Department at Eurecom in the context of a PhD thesis funded by the UCN@SOPHIA Labex.

6.2.2. *Geographically Fair In-Network Caching for Mobile Data Offloading*

Participant: Chadi Barakat

Data offloading from the cellular network to low-cost WiFi has been the subject of several research works in the last years. In-network caching has also been studied as an efficient means to further reduce cellular network traffic. In this contribution, done jointly with the Maestro project-team, we consider a scenario where mobile users can download popular contents (e.g., maps of a city, shopping information, social media, etc.) from WiFi-enabled caches deployed in an urban area. We study the optimal distribution of contents among the caches (i.e., what contents to put in each cache) to minimize users' access cost in the whole network. We argue that this optimal distribution does not necessarily provide geographic fairness, i.e., users at different locations can experience highly variable performance. In order to mitigate this problem, we propose two different cache coordination algorithms based on gossiping. These algorithms achieve geographic fairness while preserving the minimum access cost for end users. More details on this contribution can be found in [12].

6.2.3. *Virtual Service Providers (vSP)*

Participant: Damien Saucez

The ability of SOHO networks to connect to the Internet through several Internet service providers, gives high potential to enable rich cloud-based network services for enterprises. Nevertheless, it remains a huge challenge for SOHOs to leverage such multi-homing and cloud networking capabilities. For such a reason, we introduced the vSP concept (virtual Service Provider). The idea of vSP is to hide the technical complexity inherent to multi-homing and allow SOHOs to seamlessly use their cloud resources. The role of the vSP is to orchestrate traffic between the different Internet Services Providers (ISPs) in order to maximize the cloud service performance without requiring any intervention of the SOHO network administrator. This ongoing work is done in collaboration with Telecom ParisTech, Ericsson, LISPERS.net, and Cisco Systems and is presented in two papers [19], [20] and detailed in one IETF Internet-draft [19].

6.2.4. *Rules Placement Problem in OpenFlow Networks*

Participants: Xuan Nam Nguyen, Damien Saucez, Chadi Barakat and Thierry Turletti

Software-Defined Networking (SDN) abstracts low-level network functionalities to simplify network management and reduce costs. The OpenFlow protocol implements the SDN concept by abstracting network communications as flows to be processed by network elements. In OpenFlow, the high-level policies are translated into network primitives called rules that are distributed over the network. While the abstraction offered by OpenFlow allows to potentially implement any policy, it raises the new question of how to define the rules and where to place them in the network while respecting all technical and administrative requirements. We proposed a comprehensive study of the so-called OpenFlow rules placement problem with a survey of the various proposals intending to solve it [11] and developed an offline optimization framework for this problem with a polynomial time approximation in [13].

6.3. Experimental Evaluation

6.3.1. Automating ns-3 Experimentation in Multi-Host Scenarios

Participants: Alina Ludmila Quereilhac, Damien Saucez, Thierry Turletti and Walid Dabbous
ns-3 is a flexible simulator whose capabilities go beyond running purely synthetic simulations in a local desktop. Due to its ability to run unmodified Linux applications, to execute in real time mode, and to exchange traffic with live networks, ns-3 can be combined with live hosts to run distributed simulations or to transparently integrate live and simulated networks. Nevertheless, setting up ns-3 multi-host experiment scenarios might require considerable manual work and advanced system administration skills. The NEPI experiment management framework is capable of automating deployment, execution, and result collection of experiment scenarios that combine ns-3 with multiple hosts in various ways, reducing the burden of manual scenario set up. We proved that this approach can be used to seamlessly running parallel simulations on a cluster of hosts, running distributed simulation spanning multiple hosts, and integrating live and simulated networks. This work has been published in [18] and has been awarded as the best paper of the workshop.

6.3.2. DiG: Emulating Data Centers and Cloud Architectures in a Grid Network

Participants: Hardik Soni, Thierry Turletti, Damien Saucez
We are witnessing a considerable amount of research work related to data-center and cloud infrastructures but evaluations are often limited to small scale scenarios as very few researchers have access to a real infrastructure to confront their ideas to reality. We have designed an experiment automation tool, called DiG (Data-centers in the Grid), which explicitly allocates physical resources in grids to emulate data-center and cloud networks. DiG allows one to utilize grid infrastructures to evaluate research ideas pertaining to data-centers and cloud environments at massive scale and with real traffic workload. We have automated the procedure of building target network topologies while respecting effective performance capacity of available physical resources in the grid against the demand of links and hosts in the experiment. We demonstrate a showcase where DiG automatically builds a large data-center topology composed of hundreds of servers executing various Hadoop intensive workloads (see our demo abstract at IEEE NFV/SDN 2015 in [24]).

DIONYSOS Project-Team

6. New Results

6.1. Quality of Experience

Participants: Yassine Hadjadj-Aoul, Gerardo Rubino.

QoE in mobile networks. We consider in [43] an important Quality of Experience (QoE) indicator in cellular networks that is reneging of users due to impatience. We specifically consider a cell under heavy load conditions, modeled as a multiclass Processor Sharing system, and compute the reneging probability by using a fluid limit analysis. In order to enhance the user QoE, we propose a radio resource allocation control scheme that minimizes the global reneging rates. This control scheme is based on the α -fair scheduling framework and adapts the scheduler parameter depending on the traffic load. While the proposed scheme is simple, our results show that it achieves important performance gains. This work is extended in [42]. By solving the fixed point equation, we obtain a new QoE perturbation metric quantifying the impact of reneging on the performance of the system. This metric is then used to devise a new pricing scheme accounting of reneging. We specifically propose several flavors of this scheme around the idea of having a flat rate for accessing the network and an elastic price related to the level of QoE perturbation induced by communications.

In order to offer a high media quality and a good user satisfaction, the media streaming service requires that transport protocols can be adapted continuously to the network parameters. However, the diversity of terminals (e.g., tablets, smart phones, laptops) and their corresponding capabilities, mean that users' agnostic solutions are inefficient to cope with such diverse contexts. Indeed, the intrinsic characteristics and parameters of the terminals (i.e., devices) need to be taken into account on the video streaming adaptation process. In [17], we propose an adaptive video streaming solution to improve the user satisfaction factor by adapting the TCP parameters according to the user's parameters on mobile networks. The user satisfaction factor is calculated according to some metrics driven from the user's quality of experience (QoE). The work is validated through our proposal based on a new mobile agent developed on a Linux script platform and tested on different kinds of devices with different scenarios.

Learning tools. Our QoE measuring techniques (see 3.2) are based on statistical learning methods, and we have been using Random Neural Networks as our main learning tool. These are actually open queueing networks where customers have a "sign" and behave analogously as neural spiking signals. They have been proposed by Gelenbe in the 80s, and have been used in many areas since then. In [26], we published a survey about the tool, where we develop in some detail their use in supervised learning, not only for the case of interest in PSQA, our QoE measuring technology. We also discuss the use of powerful optimization methodology, first and second order techniques, that have proved to be very effective in the standard Neural Network area.

Recently, we started to explore new learning techniques. The first reason is not the search for more accurate tools, because ours are, we claim, as accurate as they can be, it is to improve robustness. The second reason is to extend our QoE measuring tools to richer contexts, mainly when we take into account time, that is, time series data. This comes from the observation that in many cases, the way people perceive quality has some "inertia" and depends on the quality perceived some minutes ago. In [66] we explored the capabilities of a recently proposed method called "Reservoir Computing (RC) with Random Static Projections" which combines two ideas, the now classic Reservoir Computing approach and Extreme Learning Machines (ELMs). In our paper, we replaced the ELMs by Radial Basis Functions (RBF) projections. We illustrated the good behavior of this variation of the original technique basically using known benchmarks.

In [67], we perform a detailed analysis of one of the main instances of the Reservoir Computing idea, called Echo State Network (ESN). This type of model has several parameters to adjust, that have an impact on the performances of the learning procedure. For instance, it has been shown that the spectral radius of the reservoir matrix (the recurrent network structure that doesn't learn during the process) is related to the accuracy and the memory capabilities of the technology. The size of the reservoir is also a parameter to adjust when configuring

an ESN for performing some specific task. One of the results of our work is the fact that the periodic or pseudo-period nature of data is also an important factor to be taken into account when designing an ESN, since it has an influence on the impact of parameters such as the previously mentioned spectral radius.

QoE and emergency management. As a by-product of our activities around QoE, we started to work on an application where, instead of evaluating the QoE of, say, a video or voice application, we wanted to evaluate the way users perceive a service not necessarily based on audio or video content. This was related to our participation to the European project QuEEN (see 8.2.2). We finished by building a platform where we test different ideas for managing an emergency situation. In our system, we include an automatic evaluator of the perceived quality of the related voice and video communications, since in the case of some catastrophes, the communications can be seriously damaged and it is critical to automatically detect the issue in order to report the problem and to take appropriate countermeasures, when possible. In [55], we describe some of the aspects of our system and of the implemented mechanisms, and we present some design problems and their solutions, together with illustrations of the capabilities of the tool.

6.2. Analytic models

Participants: Gerardo Rubino, Bruno Sericola.

Sojourn times in Markovian models. In [74], we discuss different issues related to the time a Markov chain spends in a part of its state space. This is relevant in many application areas including those interesting Dionysos, namely, performance and dependability analysis of complex systems. For instance, in dependability, the reliability of a system subject to failures and repairs of its components, is, in terms of a discrete-space model of it, the probability that it remains in the subset of operational or up states during the whole time interval $[0, t]$. In performance, the occupancy factor of some server is the probability that, in steady state, the model belongs to the subset of states where the server is busy. This book chapter reviews some past work done by the authors on this topic, and add some new insights on the properties of these sojourn times.

Queuing systems in equilibrium. In the late 70s, Leonard Kleinrock proposed a metric able to capture the tradeoff between the work done by a system and its cost, or, in terms of queuing systems, between throughput and mean response time. The new metric was called *power* and among its properties, it satisfies a nice one informally called “keep the pipe full”, specifying that the operation point of some queues (mainly the $M/M/1$ one) giving the maximal possible value to the power is when the mean backlog is 1. In [56], we took back this idea to explore what happens when we consider Jackson queuing networks. After showing that the same property holds for them and exploring other ones, we show that the power metric has some drawbacks when considering multiserver queues and networks of queues. We then propose a new metric that we called *effectiveness*, identical to power when there is a single queue with a single server, but different otherwise, that avoids these drawbacks. We analyze it and, in particular, we show that the same “keep the pipe full” holds for it.

Transient analysis of queuing systems. In a well-known book [86], today out of press, a concept of dual of a birth-and-death process is proposed, based on stochastic monotonicity. In past work [88] we showed that this concept coupled with the classical randomization or uniformization of continuous time Markov chains and lattice path combinatorics, allowed to derive analytical expressions of the transient distribution of several Markovian queuing systems. Recently, we discovered two new things: first, that this dual concept can be generalized to arbitrary systems of ordinary differential equations (ODEs) and still keep its main properties; second, that we can define a similar transformation than uniformization, that can be applied to arbitrary systems of ODEs and again, holding similar properties than the former. We respectively called pseudo-dual and pseudo-randomization the two concepts and associated methods. In [69], we presented these ideas and first results about them. We illustrated their use, and how they allow to obtain analytical expressions of transient queues’ distributions in cases where Anderson’s dual doesn’t exist (see [87]).

In [68], we present results concerning some aspects of the behavior of a queuing system observed during a fixed time period of the form $[0, t]$. The two aspects we looked at in this work are the loss process of a finite capacity model during the considered $[0, t]$, and the maximal backlog reached at a queue over the interval.

Following the classical procedure mentioned below, consisting in using uniformization to go to discrete time and then, combinatorial techniques, we develop numerical schemes to analyze both aspects of some basic queueing systems.

Network reliability. In [28], we consider the classical network design “Capacitated m -Ring Star Problem” (CmRSP), where we look for m rings connecting two nodes in a network at minimum cost. We add to this model the fact that links can fail, and propose a new paradigm that we call “Capacitated m -Ring Star Problem with Diameter Constrained Reliability ” (in short, CmRSP-DCR), where we look again for a minimal cost spanning graph of the set of nodes in the network that connects the selected source and terminal, *while satisfying a Diameter Constrained Reliability (DCR) condition*. The DCR is the probability that the two nodes can communicate by means of paths having lengths bounded by some fixed value d . We prove that this problem is NP-hard, and we propose a GRASP-based approach to solve it.

Fluid models. In [19] we study congestion periods in a finite fluid buffer when the net input rate depends upon a recurrent Markov process; congestion occurs when the buffer content is equal to the buffer capacity. We consider the duration of congestion periods as well as the associated volume of lost information. We derive their distributions in a typical stationary busy period of the buffer. Our goal is to compute the exact expression of the loss probability in the system, which is usually approximated by the probability that the occupancy of the infinite buffer is greater than the buffer capacity under consideration. Moreover, by using general results of the theory of Markovian arrival processes, we show that the duration of congestion and the volume of lost information have phase-type distributions.

6.3. Performance Evaluation of Distributed Systems

Participants: Bruno Sericola, Yann Busnel, Pierre L’Ecuyer.

Detection of distributed deny of service attacks. A Deny of Service (DoS) attack tries to progressively take down an Internet resource by flooding this resource with more requests than it is capable to handle. A Distributed Deny of Service (DDoS) attack is a DoS attack triggered by thousands of machines that have been infected by a malicious software, with as immediate consequence the total shut down of targeted web resources (e.g., e-commerce websites). A solution to detect and to mitigate DDoS attacks is to monitor network traffic at routers and to look for highly frequent signatures that might suggest ongoing attacks. A recent strategy followed by the attackers is to hide their massive flow of requests over a multitude of routes, so that locally, these flows do not appear as frequent, while globally they represent a significant portion of the network traffic. The term “iceberg” has been recently introduced to describe such an attack as only a very small part of the iceberg can be observed from each single router. The approach adopted to defend against such new attacks is to rely on multiple routers that locally monitor their network traffic, and upon detection of potential icebergs, inform a monitoring server that aggregates all the monitored information to accurately detect icebergs [36]. Now to prevent the server from being overloaded by all the monitored information, routers continuously keep track of the c (among n) most recent high flows (modeled as items) prior to sending them to the server, and throw away all the items that appear with a small probability. Parameter c is dimensioned so that the frequency at which all the routers send their c last frequent items is low enough to enable the server to aggregate all of them and to trigger a DDoS alarm when needed. This amounts to compute the time needed to collect c distinct items among n frequent ones. A thorough analysis of the time needed to collect c distinct items appears in [12], [11].

Stream Processing Systems. Stream processing systems are today gaining momentum as tools to perform analytics on continuous data streams. Their ability to produce analysis results with sub-second latencies, coupled with their scalability, makes them the preferred choice for many big data companies.

A stream processing application is commonly modeled as a direct acyclic graph where data operators, represented by nodes, are interconnected by streams of tuples containing data to be analyzed, the directed edges (the arcs). Scalability is usually attained at the deployment phase where each data operator can be parallelized using multiple instances, each of which will handle a subset of the tuples conveyed by the operators’ ingoing stream. Balancing the load among the instances of a parallel operator is important as it yields to better resource

utilization and thus larger throughputs and reduced tuple processing latencies. We have proposed a new key grouping technique targeted toward applications working on input streams characterized by a skewed value distribution [53]. Our solution is based on the observation that when the values used to perform the grouping have skewed frequencies, the few most frequent values (the *heavy hitters*) drive the load distribution, while the remaining largest fraction of the values (the *sparse items*) appear so rarely in the stream that the relative impact of each of them on the global load balance is negligible. We have shown, through a theoretical analysis, that our solution provides on average near-optimal mappings using sub-linear spaces in the number of tuples read from the input stream in the learning phase and the support (value domain) of the tuples. In particular this analysis presents new results regarding the expected error made on the estimation of the frequency of heavy hitters.

Randomized Message-Passing Test-and-Set. In [37], we have presented a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes, returns yes to a unique process and returns no to all the others. Recently, many advances in implementing Test&Set objects have been achieved. However, all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p \leq n$ of processes where n is the total number of processes in the system. It has an expected individual step complexity in $O(\log p)$ against an oblivious adversary, and an expected individual message complexity in $O(n)$. The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

Population Protocol Model. The population protocol model, introduced by Angluin and his colleagues in 2006, provides theoretical foundations for analyzing global properties emerging from pairwise interactions among a large number of anonymous agents. In the population protocol model, agents are modeled as identical and finite state machines, i.e each agent can be in a finite number of states while waiting to execute a transition. When two agents interact, they communicate their local state, and can move from one state to another according to a transition function. The ultimate goal of population protocols is for all the agents to converge to the same value. Examples of systems whose behavior can be modeled by population protocols range from molecule interactions of a chemical process to sensor networks in which agents, which are small devices embedded for instance in animals, interact each time two animals are in the same radio range.

In this work, we focus on a quite important related question. Namely, is there a population protocol that exactly counts the difference κ between the number of agents that initially set their state to A and the one that initially set it to B , and can it be solved in an efficient way, that is with the guarantee that each agent should converge to the exact value of κ after having triggered a sub-linear number of interactions in the size of the system [49]? We answer this question by the affirmative by presenting a $O(n^{3/2})$ -state population protocol that allows each agent to converge to the exact solution by interacting no more than $O(\log n)$ times. The proposed protocol is very simple (as is true for most known population protocols), but is general enough to be used to solve different types of tasks.

Call centers. We develop research activities around the analysis and design of call centers, from a performance perspective. The effective management of call centers is a challenging task mainly because managers are consistently facing considerable uncertainty. Among important sources of uncertainty are call arrival rates which are typically time-varying, stochastic, dependent across time periods and across call types, and often affected by external events. Accurately modeling and forecasting future call arrival volumes is a complicated issue which is critical for making important operational decisions, such as staffing and scheduling, in the call center. In [20] we review the existing literature on modeling and forecasting call arrivals. We also develop in [58] customer delay predictors for multi-skill call centers that take as inputs the queueing state upon arrival and the waiting time of the last customer served. Barely any predictor currently exists for the multi-skill case. We introduce two new predictors that use cubic regression splines and artificial neural networks, respectively, and whose parameters are optimized (or learned) from observation data obtained by simulation.

6.4. Wireless Networks

Participants: Osama Arouk, Btissam Er-Rahmadi, Adlen Ksentini, Meriem Bouzouita, Pantelis Frangoudis, Yassine Hadjadj-Aoul, Gerardo Rubino.

We are continuing our activities around wireless and mobile networks, by focusing more on leveraging the current mobile and wireless architecture toward building the 5G systems.

LTE improvements. One of the 5G objectives is to support a high number of devices. This not only concerns User Equipment (UE) devices, but also other devices such as sensors and actuators (known also as Internet of Things (IoT)). Sensor and actuator devices communicate generally with a remote server in an automatic way, without any human intervention. This type of communication is known as Machine to Machine (M2M) communication, or Machine Type Communication (MTC). The corresponding traffic is known by its intensity and impact on increasing congestion in both main parts of 4G networks, the Radio Access Network (RAN) and the Core Network. To improve the current LTE system to support MTC, we did several contributions. We proposed in [51] an important enhancement to the Group Paging (GP) mechanism, which is responsible for relaying requests to sensors, in order to gather data. After modeling analytically the GP procedure, we proposed a mechanism that, instead of paging all MTC devices in the same period, calculates the appropriate number of MTCs that reduces the collision probability as well as increases the success probability. In [52], we modeled the Radio Access Channel (RACH) procedure when the MTC devices are activated in a highly synchronized manner during a certain period (synchronized traffic), which is represented by a Beta distribution. The proposed model estimates for each period the exact number of MTC devices that may win the contention.

To control the Random Access Network (RAN) overload and alleviate the access network congestion, 3GPP developed the Access Class Barring (ACB) procedure that depends on an access probability called the ACB factor, without proposing a procedure for calculating such probability. In [72], we have proposed a fluid-based random access model for M2M communications, which was used to determine dynamically the value of the ACB factor that avoids system overload and the radio resources' underutilization at the same time. We proposed in [60] a novel implementation of the ACB mechanism in the context of multiple M2M traffic classes. Based on a scheduling algorithm, we have applied a PID controller to adjust dynamically multiple ACB factors related to each class category, guaranteeing a number of devices around an optimal value that maximizes the Random Access (RA) success probability. In [61], we first present a simple fluid model of MTC devices' random access. This model is then used to derive a novel adaptive regulator of the ACB factor, somehow in contrast with previous existing contributions which generally rely on heuristics. The main advantages of the proposed approach are twofold. First, the proposal is fully compliant with the standard while it reduces significantly the computation and the signaling overheads. Second, it provides an efficient mean to regulate adaptively the ACB factor as it guarantees having an optimal number of MTC devices accessing concurrently to the RAN. The obtained results based on simulations show clearly the robustness of the proposed approach, and its superiority compared to existing proposals.

Another important objective of 5G mobile networks is to accommodate a diverse and ever-increasing number of user equipments (UEs). Coping with the massive signaling overhead expected from UEs is an important hurdle to tackle so as to achieve this objective. In [38], we devised an efficient tracking area list management framework that aims for finding optimal distributions of tracking areas (TAs) in the form of TA lists (TALs) and assigning them to UEs. The objective is to minimize two conflicting metrics: paging overhead and tracking area update (TAU) overhead. We used bargaining games to find the Pareto optimal solution that satisfies both objectives.

WiFi networks improvements. It is well established that WiFi is complementing LTE connections to ensure, wirelessly, high data rate. One idea to improve WiFi towards high data rates is to multiple users' transmissions on both directions, i.e. on the Down Link (DL) and the Up Link (UL). In [50] we devised a novel solution to enhance the TXOP Sharing mechanism, introduced in the 802.11ac amendment, to achieve efficient Down-Link Multi-User Multiple-Input Multiple-Output (DL-MU-MIMO) transmission. First, we give new definitions about both events of successful and failed DL-MU-MIMO transmission. Then, we devise a revised

Backoff procedure for the primary Access Category (AC). In [40] we proposed a novel 802.11ax MAC protocol aiming at reducing the elapsed time in managing the establishment of an UL-MU communication, thus enhancing considerably the system's performance.

On the other hand, the volume of mobile multimedia traffic is fast-growing, challenging the radio and backhaul network infrastructure and calling for alternative content dissemination schemes. To improve user experience and reduce infrastructure load, we exploit implicit social relationships among users and take into account content popularity, proposing push-based prefetching mechanisms which take advantage of the caching and mobile ad hoc networking capabilities of user devices. We use, in [65], bloom filters as summaries of user caches, and design mechanisms to estimate the social distance between users and the popularity of content items, which drive our algorithms. Our simulation-based evaluation shows that our scheme brings caching performance improvements in an order of 10% in terms of absolute cache hit ratio in most of the cases studied, and from 3% to 82% in terms of normalized cache hit ratio gain.

Network selection. With the explosion of mobile data traffic, the Fixed and Mobile Converged (FMC) network are being heavily required. Mobile devices have the capability of connecting to different access networks in the FMC architecture simultaneously. Access network selection becomes an issue when mobile devices are under coverage of different access networks, since a bad selection may lead to network congestion and degrade the QoE of users. In order to address this problem, we model and analyze, in [62] and [63], the interface selection procedure using control theory in the FMC architecture. Based on our model, we designed a controller which can send to mobile devices a network selection command calculated instantly for the access network selection. In [29], we investigated network decentralization in conjunction with the selective IP traffic offload approaches to handle the increased data traffic. We first devised different approaches based on a per-destination-domain-name basis, which offer operators a fine-grained control to determine whether a new IP connection should be offloaded or accommodated via the core network.

Energy efficiency. Due to the ever-growing gap between battery lifetime and hardware/software complexity in addition to application's computing power needs, the energy saving issue becomes crucial. In this context, we proposed, in [13], an end-to-end study of video decoding on different architectures. The study was achieved thanks to a two steps methodology: (1) a comprehensive characterization and evaluation of the performance and the energy consumption of video decoding, (2) an accurate high level energy model based on the characterization step. In [24], we proposed to apply data fragmentation, in slotted CSMA/CA, in a way to allow improving the bandwidth occupation while reducing the latency. We proposed to introduce a network allocation vector (NAV) in the fragmentation mechanism to reduce energy consumption in IEEE 802.15.4. A Markov chain-based analytical model of the fragmentation mechanism was given as well as an analytical model of the energy consumption using a NAV. The analytical results show that the fragmentation technique improves at the same time the throughput, the access delay and the bandwidth occupation. They also show that the NAV mechanism reduces energy consumption when applying the fragmentation technique in slotted CSMA-CA for IEEE 802.15.4.

6.5. Future networks and architectures

Participants: Adlen Ksentini, Yassine Hadjadj-Aoul, Jean-Michel Sanner.

SDN. We started an activity on Software Defined Networking (SDN), a recent idea proposed to handle network management problems. SDN are becoming an important issue with the ever-increasing network complexity. They are proposed as an alternative to the current architecture of the Internet, which cannot meet the supported services requirements such as Quality of Service/Experience (Qos/QoE), security and energy consumption. We particularly address the scalability issue by proposing in [70] an automated hierarchical controller-based architecture handling the whole control chain.

Mobile cloud. One of the 5G-architecture visions considers the usage of cloud to ease mobile networks evolution towards more flexibility and elasticity for handling resources; building the concept of carrier cloud. Software Defined Networking (SDN) and Network Function Virtualization (NFV) represent the key enabler of carrier cloud. In [57], we addressed the problem of Virtual Network Function (VNF) placement in the carrier

cloud. Indeed, we proposed a placement solution that has two main design goals: i) minimizing path between users and their respective data anchor gateways and ii) optimizing their sessions' mobility. The two design goals effectively represent two conflicting objectives that we deal with considering the mobility features and service usage behavioral patterns of mobile users, in addition to the mobile operators' cost in terms of the total number of instantiated VNFs to build a Virtual Network Infrastructure (VNI). We modeled this problem using an optimization formulation having these conflicting objectives, and then used Bargaining Game to find the Pareto optimal solution. We are continuing our improvement to the Follow Me Cloud (FMC), which was devised by our team conjointly with NEC labs. In [33], we proposed a FMC architecture that relies on PMIPv6 to handle mobility, and SDN to update the flow table of the anchor routers when a service has moved from one Data Center to another. In [10] and [32], we addressed the challenge of flow table scalability problem, which may arise in FMC to high number of mobile users. To this aim, we proposed a two-level hierarchical SDN controllers architecture in order to distribute the SDN/OpenFlow control plane. Another objective of 5G is to reduce network latency to 1ms, which will ease computation offloading. Thus, it will be possible to run applications on UE device, even if the latter has low computation capability, by offloading part of the code to a remote server. In [44], we were interested on studying the opportunities to offload part of one of the well known game engine in the literature, i.e. Unity 3D. We built a data set representing the CPU-GPU use of several games; allowing us to understand which modules might be offloaded to a remote server in the Mobile Cloud.

6.6. Network Economics

Participant: Bruno Tuffin.

The general field of network economics, analyzing the relationships between all acts of the digital economy, has been an important subject for years in the team. The whole problem of network economics, from theory to practice, describing all issues and challenges, is described in our book "Telecommunication Network Economics From Theory to Applications" (P. Maillé and B. Tuffin, Cambridge U. Press, 2014).

Network neutrality. Among the topics we have particularly focused on, the network neutrality debate was a major concern in 2015. In [23], [80], [83] we recall the debate and highlight the fact that neutrality principles can be bypassed in many ways without violating the rules currently evoked in the debate. For example via Content Delivery Networks (CDNs), which deliver content on behalf of content providers for a fee, or via search engines, which can hinder competition and innovation by affecting the visibility and accessibility of content. In [23], we challenge the definition of net neutrality as it is generally discussed. Our goal there is to initiate a relevant debate for net neutrality in an increasingly complex Internet ecosystem, and to provide examples of possible neutrality rules for different levels of the delivery chain, this level separation being inspired by the OSI layer model.

As particular ways to bypass the current neutrality principles, we have particularly focused on CDNs. We for example investigate in [47] the impact of decisions made by a CDN willing to maximize its revenue through the management of cache servers. Based on a model with two network providers, we highlight that revenue-oriented management policies can affect the user-perceived quality of experience, impacting the competition among network access providers in favor of the largest one. Since this contradicts the principle underpinning network neutrality?although not with the technical net neutrality rules?we discuss the necessity to regulate CDN activity. Also, one of the main argument toward neutrality being that it favors innovation, we study in [46] the impact of CDNs' activity on other actors of the supply chain. Our findings indicate that vertically integrating a CDN helps Internet Service Providers (ISPs) collect fees from Content Providers (CPs), hence circumventing the interdiction of side payments coming from net neutrality rules. However, this outcome is socially much better in terms of user quality and innovation fostering than having separate actors providing the access and CDN services: in the latter case double marginalization (both ISP and CDN trying to get some value from the supply chain) leads to suboptimal investments in CDN storage capacities and higher prices for CPs, resulting in reduced innovation.

Another model we have developed is for understanding the behavior of some big providers actually paying side payment to ISPs while still officially in favor of neutrality. To better understand this strategical behavior, we have presented a simple model in [59] providing some insight on whether or not paying side payments for an incumbent provider is a way to create barriers to entry for competitors. It also investigates the economic consequences on all actors: incumbent and new entrant content providers, users, and the Internet Service Provider. It then describes how the side payment can be determined as a Nash bargaining solution.

Pricing access networks. Access networks in a competitive context has been a topic of research for a while. In the Internet, the data charging scheme has usually been flat rate. But more recently, especially for mobile data traffic, we have seen more diversity in the pricing offers, such as volume-based ones or cap-based ones. We study in [48] the behavior of heterogeneous users facing two offers: a volume-based one and a flat-rate one. On top of that selection, we investigate 1) the relevance for an ISP to propose the two types of offers, and optimize the corresponding prices, and 2) the existence of a solution to the pricing game when the offers come from competing providers.

Sponsored auctions. Advertisement in dedicated webpage spaces or in search engines sponsored slots is usually sold using auctions, with a payment rule that is either per view or per click. But advertisers can be both sensitive to being viewed (brand awareness effect) and being clicked (conversion into sales). In [84], we generalize the auction mechanism by including both pricing components: the displayed advertisers are charged when their ad is displayed, and pay an additional price if the ad is clicked. Applying the results for Vickrey-Clarke-Groves (VCG) auctions, we show how to compute payments to ensure incentive compatibility from advertisers as well as maximize the total value of the advertisement slot(s). We provide tight upper bounds for the loss of efficiency due to applying only pay-per-click (or pay-per-view) pricing instead of our scheme. Those bounds depend on the joint distribution of advertisement visibility and population likelihood to click on ads, and can help identify situations where our mechanism yields significant improvements. We also describe how the commonly used generalized second price (GSP) auction can be extended to this context.

6.7. Monte Carlo

Participants: Pierre L'Ecuyer, Gerardo Rubino, Bruno Tuffin.

We maintain a research activity in different areas related to dependability, performability and vulnerability analysis of communication systems, using both the Monte Carlo and the Quasi-Monte Carlo approaches to evaluate the relevant metrics. Monte Carlo (and Quasi-Monte Carlo) methods often represent the only tool able to solve complex problems of these types. However, when the events of interest are rare, simulation requires a special attention, to accelerate the occurrence of the event and get unbiased estimators of the event of interest with a sufficiently small relative variance. This is the main problem in the area. Dionysos' work focuses then on dealing with the rare event situation. For example, [39] presents an exponential tilting method for exact simulation from the truncated multivariate student-t distribution in high dimensions as an alternative to approximate Markov Chain Monte Carlo sampling.

A non-negligible part of our activity on the application of rare event simulation was about the evaluation of static network reliability models. Our paper [16] focuses on a technique known as Recursive Variance Reduction (RVR) which approaches the unreliability by recursively reducing the graph from the random choice of the first working link on selected cuts. This previously known method is shown to not verify the bounded relative error (BRE) property as reliability of individual links goes to one, i.e., the estimator is not robust in general to high reliability of links. We then propose to use the decomposition ideas of the RVR estimator in conjunction with the IS technique. Two new estimators are presented in the paper: the first one, called Balanced Recursive Decomposition estimator, chooses the first working link on cuts uniformly, while the second, called Zero-Variance Approximation Recursive Decomposition estimator, combines RVR and our zero-variance IS approximation. We show that in both cases BRE property is verified and, moreover, that a vanishing relative error (VRE) property can be obtained for the Zero-Variance Approximation RVR under specific sufficient conditions. A numerical illustration of the power of the methods is provided on several benchmark networks. In [54], we explore the use of the same powerful RVR idea, but applied in a very general

context, where the system is model by a monotone structure function. In the paper, we illustrate the approach with a very widely used model, a series of k -out-of- m modules.

In a static network reliability model one typically assumes that the failures of the components of the network are independent. This simplifying assumption makes it possible to estimate the network reliability efficiently via specialized Monte Carlo algorithms. Hence, a natural question to consider is whether this independence assumption can be relaxed, while still attaining an elegant and tractable model that permits an efficient Monte Carlo algorithm for unreliability estimation. In [14] we provide one possible answer by considering a static network reliability model with dependent link failures, based on a Marshall-Olkin copula, which models the dependence via shocks that take down subsets of components at exponential times, and propose a collection of adapted versions of permutation Monte Carlo (PMC, a conditional Monte Carlo method), its refinement called the turnip method, and generalized splitting (GS) methods, to estimate very small unreliabilities accurately under this model. The PMC and turnip estimators have bounded relative error when the network topology is fixed while the link failure probabilities converge to 0, whereas GS does not have this property. But when the size of the network (or the number of shocks) increases, PMC and turnip eventually fail, whereas GS works nicely (empirically) for very large networks, with over 5000 shocks in our examples. In [41] we focus on a method proposed by Fishman making use of bounds on the structure function describing in terms of configurations of (independent) link states if the considered nodes are connected. The bounds are based on the computation of (independent) mincuts disconnecting the set of nodes and (independent) minpaths ensuring that they are connected. We analyze here the robustness of the method when the unreliability of links goes to zero. We show that the conditions provided by Fishman are based on a bound and are therefore only sufficient, and provide more insight and examples on the behavior of the method.

PMC is an effective way of estimating the unreliability of a static network when this unreliability is very small and the network is not too large. We generalize the method in [31] to cover a wider range of applications, in which an estimation problem can be reframed in terms of the hitting time of a given set of states by a continuous-time Markov chain. The estimator is then defined as a function of the sample path of the underlying discrete time chain only, via Conditional Monte Carlo. We prove that the method gives bounded relative error for rare event probability estimation in certain settings. We show how it can be used to estimate the cumulative distribution function, or the density, or some moment of the hitting time. We provide examples for which the method can be applied and we give numerical illustrations.

Another family of models of interest in the group are the highly reliable Markovian systems, where a Markov chain models the evolution of a multicomponent system with failures and repairs of its components. In [27] we explore a new approach in the context of these models, and in the rare event case, called Conditional Monte Carlo with Intermediate Estimations (CMIE). The target are models with complex structures, where it is hard to design a good *importance function* dealing to good Importance Sampling schemes. The paper shows that the method belongs to the variance reduction family, and some examples illustrate its performances. It can be seen as a generalization of the class of splitting simulation procedures.

Finally, in Quasi-Monte Carlo (QMC), we reviewed in [64] the recent development on array-RQMC, a randomized quasi-Monte Carlo method for we had developed estimating the state distribution at each step of a Markov chain with totally ordered (discrete or continuous) state space. It can be used in particular to obtain a low-variance unbiased estimator of the expected total cost up to some random stopping time, when state-dependent costs are paid at each step. In [21], a combination of sequential MC with RQMC to accelerate convergence proposed by Gerber and Chopin is compared with our array-RQMC.

But simulation requires the use of pseudo-random generators. In [45] we provide a review of the state of the art on the design and implementation of random number generators (RNGs) for simulation, on both sequential and parallel computing environments. A general review of pseudo-random and quasi-random number generation is also provided in [73]. A tool for the generation of rank-1 lattice rules is described in [22].

DIVERSE Project-Team

7. New Results

7.1. Results on Software Language Engineering

7.1.1. Modular and Reusable Development of DSLs

Domain-Specific Languages (DSLs) are now developed for a wide variety of domains to address specific concerns in the development of complex systems. When engineering new DSLs, it is likely that previous efforts spent on the development of other languages could be leveraged, especially when their domains overlap. However, legacy DSLs may not fit exactly the end user requirements and thus require further extension, restriction, or specialization. While current language workbenches provide import mechanisms, they usually lack an explicit support for such customizations of imported artifacts. We propose an approach for building DSLs by safely assembling and customizing legacy DSLs artifacts. This approach is based on typing relations that provide a reasoning layer for manipulating DSLs while ensuring type safety. On top of this reasoning layer, we provide an algebra of operators for extending, restricting, and assembling separate DSL artifacts. We implemented the typing relations and algebra into the Melange meta-language [30], [29], [73].

7.1.2. Executable Domain-Specific Modeling Languages (xDSMLs)

Executable Domain-Specific Modeling Languages (xDSMLs) open many possibilities for performing early verification and validation (V&V) of systems. Dynamic V&V approaches rely on execution traces, which represent the evolution of models during their execution. In order to construct traces, generic trace metamodels can be used. Yet, regarding trace manipulations, they lack both efficiency because of their sequential structure, and usability because of their gap to the xDSML. We contributed a generative approach that defines a rich and domain-specific trace metamodel enabling the construction of execution traces for models conforming to a given xDSML [24]. We also contributed a partly generic omniscient debugger supported by generated domain-specific trace management facilities [49].

The emergence of modern concurrent systems calls for xDSMLs where concurrency is of paramount importance. Such xDSMLs are intended to propose constructs with rich concurrency semantics, which allow system designers to precisely define and analyze system behaviors. In [34], we introduce a concurrent executable metamodeling approach, which supports a modular definition of the execution semantics, including the concurrency model, the semantic rules, and a well-defined and expressive communication protocol between them. In [28], we present MoCCML, a dedicated meta-language for formally specifying the concurrency concern within the definition of a DSL. The concurrency constraints can reflect the knowledge in a particular domain, but also the constraints of a particular platform. MoCCML comes with a complete language workbench to help a DSL designer in the definition of the concurrency directly within the concepts of the DSL itself, and a generic workbench to simulate and analyze any model conforming to this DSL. MoCCML is illustrated on the definition of an lightweight extension of SDF (SynchronousData Flow).

7.1.3. Globalization of Domain-Specific Modeling Languages

The development of modern complex software-intensive systems often involves the use of multiple DSMLs that capture different system aspects. Supporting coordinated use of DSMLs leads to what we call the globalization of modeling languages, that is, the use of multiple modeling languages to support coordinated development of diverse aspects of a system.

In a book published in 2015 [66], a number of articles describe the vision and the way globalized DSMLs currently assist integrated DSML support teams working on systems that span many domains and concerns to determine how their work on a particular aspect influences work on other aspects. Globalized DSMLs offer support for communicating relevant information, and for coordinating development activities and associated technologies within and across teams, in addition to providing support for imposing control over development artifacts produced by multiple teams. DSMLs can be used to support socio-technical coordination by providing the means for stakeholders to bridge the gap between how they perceive a problem and its solution, and the programming technologies used to implement a solution. They also support coordination of work across multiple teams. DSMLs developed in an independent manner to meet the specific needs of domain experts have an associated framework that regulates interactions needed to support collaboration and work coordination across different system domains. The book includes [63], [65], [64], [62] with authors from the DIVERSE team.

In [43], we propose a Behavioral Coordination Operator Language (B-COOL) to reify coordination patterns between specific domains by using coordination operators between the Domain-Specific Modeling Languages used in these domains. Those operators are then used to automate the coordination of models conforming to these languages. We illustrate the use of B-COOL with the definition of coordination operators between timed finite state machines and activity diagrams.

The GEMOC Studio (<http://gemoc.org/studio>) is an eclipse package that contains components for building and composing executable Domain-Specific Modeling Languages (DSMLs). The GEMOC Studio complements Melange to formally define in a modular way the concurrency model of executable DSMLs, and provides analysis and coordination facilities based on the concurrency model. It also integrates all the contributions presented in this document related to model execution, animation, debugging and trace management. The GEMOC studio has been the overall winner of the transformation tool contest 2015 on Model Execution [52].

7.1.4. An analysis of metamodeling practices for MOF and OCL

The definition of a metamodel that precisely captures domain knowledge for effective know-how capitalization is a challenging task. A major obstacle for domain experts who want to build a metamodel is that they must master two radically different languages: an object-oriented, MOF-compliant, modeling language to capture the domain structure and first order logic (the Object Constraint Language) for the definition of well-formedness rules. However, there are no guidelines to assist the conjunct usage of both paradigms, and few tools support it. Consequently, we observe that most metamodels have only an object-oriented domain structure, leading to inaccurate metamodels. In [21], we perform the first empirical study, which analyzes the current state of practice in metamodels that actually use logical expressions to constrain the structure. We analyze 33 metamodels including 995 rules coming from industry, academia and the Object Management Group, to understand how metamodelers articulate both languages. We implement a set of metrics in the OCLMetrics tool to evaluate the complexity of both parts, as well as the coupling between both. We observe that all metamodels tend to have a small, core subset of concepts, which are constrained by most of the rules, in general the rules are loosely coupled to the structure and we identify the set of OCL constructs actually used in rules.

7.1.5. Model Slicers

Among model comprehension tools, model slicers are tools that extract a subset of model elements, for a specific purpose. We propose the Kompren language to model and generate model slicers for any DSL (*e.g.* modeling for software development or for civil engineering) and for different purposes (*e.g.* monitoring and model comprehension). We detail the semantics of the Kompren language and of the model slicer generator. This provides a set of expected properties about the slices that are extracted by the different forms of the slicer [18]. We show how the use of Kompren, a domain-specific language for defining model slicers, can ease the development of such interactive visualization features [19].

In Model Driven Development (MDD), it is important to ensure that a model conforms to the invariants defined in the metamodel. General-purpose rigorous analysis tools that check invariants are likely to perform the analysis over the entire metamodel and model. Since modern day software is exceedingly complex, the

size of the model together with the metamodel can be very large. Consequently, invariant checking can take a very long time. To this end, we introduce model slicing within the invariant checking process, and use a slicing technique to reduce the size of the inputs in order to make invariant checking of large models feasible with existing tools [22], [42].

7.1.6. Bridging the gap between scientific models and engineering models with MDE

The complex problems that computational science addresses are more and more benefiting from the progress of computing facilities (e.g., simulators, libraries, accessible languages). Nevertheless, the actual solutions call for several improvements. Among those, we address in [25] the needs for leveraging on knowledge and expertise by focusing on Domain-Specific Modeling Languages application. In this vision paper we illustrate, through concrete experiments, how the last DSML research help getting closer the problem and implementation spaces.

Various disciplines use models for different purposes. While engineering models, including software engineering models, are often developed to guide the construction of a non-existent system, scientific models, in contrast, are created to better understand a natural phenomenon (i.e., an already existing system). An engineering model may incorporate scientific models to build a system. Both engineering and scientific models have been used to support sustainability, but largely in a loosely-coupled fashion, independently developed and maintained from each other. Due to the inherent complex nature of sustainability that must balance trade-offs between social, environmental, and economic concerns, modeling challenges abound for both the scientific and engineering disciplines. In [72] we propose a vision that synergistically combines engineering and scientific models to enable broader engagement of society for addressing sustainability concerns, informed decision-making based on more accessible scientific models and data, and automated feed-back to the engineering models to support dynamic adaptation of sustainability systems. To support this vision, we identify a number of challenges to be addressed with particular emphasis on the socio-technical benefits of modeling.

As first experiments, we presented at the Inria-Industry meeting 2015 on energy transition and EclipseCon 2015, an approach to develop smart cyber physical systems in charge of managing the production, distribution and consumption of energies (e.g., water, electricity). The main objective is to enable a broader engagement of society, while supporting a more informed decision-making, possibly automatically, on the development and run-time adaptation of sustainability systems (e.g., smart grid, home automation, smart cities). We illustrate this approach through a system that allows farmers to simulate and optimize their water consumption by combining the model of a farming system together with agronomical models (e.g., vegetable and animal lifecycle) and open data (e.g., climate series). To do so, we use Model Driven Engineering (MDE) and Domain Specific Languages (DSL) to develop such systems driven by scientific models that define the context (e.g., environment, social and economy), and model experiencing environments to engage general public and policy makers.

7.2. Results on Variability Modeling and Engineering

7.2.1. Reverse engineering variability

We have developed automated techniques and a comprehensive environment for synthesizing feature models from various kinds of artefacts (e.g. propositional formula, dependency graph, FMs or product comparison matrices). Specifically we have elaborated a support (through ranking lists, clusters, and logical heuristics) for choosing a sound and meaningful hierarchy [93]. We have performed an empirical evaluation on hundreds of feature models, coming from the SPLOT repository and Wikipedia [92]. We have showed that a hybrid approach mixing logical and ontological techniques outperforms state-of-the-art solutions (to appear in Empirical Software Engineering journal in 2015 [20]). We have also considered numerical information and feature *attributes* so that we are now capable of synthesizing attributed feature models from product descriptions [51].

Besides, we have developed techniques for reverse engineering variability in generators and configurators (e.g., video generators) [50]. We have identified new research directions for protecting variability [44] mainly due to the fact reverse engineering techniques (previously presented) are effective .

7.2.2. Product comparison matrices

Product Comparison Matrices (PCMs) constitute a rich source of data for comparing a set of related and competing products over numerous features. PCMs can be seen as a formalism for modeling a family of products, including variability information. Despite their apparent simplicity, PCMs contain heterogeneous, ambiguous, uncontrolled and partial information that hinders their efficient exploitations. We have formalized PCMs through model-based automated techniques and developed additional tooling to support the edition and re-engineering of PCMs [94]. 20 participants used our editor to evaluate our PCM metamodel and automated transformations. The empirical results over 75 PCMs from Wikipedia show that (1) a significant proportion of the formalization of PCMs can be automated: 93.11% of the 30061 cells are correctly formalized; (2) the rest of the formalization can be realized by using the editor and mapping cells to existing concepts of the metamodel. The ASE'2014 paper opens avenues for engaging a community in the mining, re-engineering, edition, and exploitation of PCMs that now abound on the Internet. We have launched an open, collaborative initiative towards this direction <https://opencompare.org/>

Another axis is the mining of PCMs since (1) the manual elaboration of PCMs has limitations (2) numerous sources of information can be combined and are amenable to PCMs. We have developed MatrixMiner a tool for automatically synthesizing PCMs from a set of product descriptions written in natural language [46]. MatrixMiner is capable of identifying and organizing features and values in a PCM despite the informality and absence of structure in the textual descriptions of products. More information is available online: <https://matrix-miner.variability.io/>

7.3. Results on Heterogeneous and dynamic software architectures

7.3.1. Resource Monitoring and Reservation in Heterogeneous and dynamic software architectures

Software systems are more pervasive than ever nowadays. Occasionally, applications run on top of resource-constrained devices where efficient resource management is required; hence, they must be capable of coping with such limitations. However, applications require support from the runtime environment to properly deal with resource limitations. This thesis addresses the problem of supporting resource-aware programming in execution environments. In particular, it aims at offering efficient support for collecting data about the consumption of computational resources (e.g., CPU, memory), as well as efficient mechanisms to reserve resources for specific applications. In existing solutions we find two important drawbacks. First, they impose performance overhead on the execution of applications. Second, creating resource management tools for these abstractions is still a daunting task. The outcomes of this work [12] are three contributions:

- An optimistic resource monitoring framework that reduces the cost of collecting resource consumption data.
- A methodology to select components' bindings at deployment time in order to perform resource reservation.
- A language to build customized memory profilers that can be used both during applications' development, and also in a production environment.

7.3.2. Dynamic Reasoning on Heterogeneous and dynamic software architectures

Multi-Objective Evolutionary Algorithms (MOEAs) have been successfully used to optimize various domains such as finance, science, engineering, logistics and software engineering. Nevertheless, MOEAs are still very complex to apply and require detailed knowledge about problem encoding and mutation operators to obtain an effective implementation. Software engineering paradigms such as domain-driven design aim to tackle this complexity by allowing domain experts to focus on domain logic over technical details. Similarly, in order to handle MOEA complexity, we propose an approach, using model-driven software engineering (MDE) techniques, to define fitness functions and mutation operators without MOEA encoding knowledge. Integrated into an open source modelling framework, our approach can significantly simplify development and maintenance of multi-objective optimizations. By leveraging modeling methods, our approach allows reusable

optimizations and seamlessly connects MOEA and MDE paradigms. We evaluate our approach on a cloud case study and show its suitability in terms of i) complexity to implement an MOO problem, ii) complexity to adapt (maintain) this implementation caused by changes in the domain model and/or optimization goals, and iii) show that the efficiency and effectiveness of our approach [56] remains comparable to ad-hoc implementations.

7.3.3. A Precise Metamodel for Open Cloud Computing Interface

Open Cloud Computing Interface (OCCI) proposes one of the first widely accepted, community-based, open standards for managing any kinds of cloud resources. But as it is specified in natural language, OCCI is imprecise, ambiguous, incomplete, and needs a precise definition of its core concepts. Indeed, the OCCI Core Model has conceptual drawbacks: an imprecise semantics of its type classification system, a nonextensible data type system for OCCI attributes, a vague and limited extension concept and the absence of a configuration concept. To tackle these issues, this work proposes a precise metamodel for OCCI. This metamodel defines rigorously the static semantics of the OCCI core concepts, of a precise type classification system, of an extensible data type system, and of both extension and configuration concepts. This metamodel is based on the Eclipse Modeling Framework (EMF), its structure is encoded with Ecore and its static semantics is rigorously defined with Object Constraint Language (OCL). As a consequence, this metamodel provides a concrete language to precisely define and exchange OCCI models. The validation of our metamodel is done on the first worldwide dataset of OCCI extensions already published in the literature, and addressing inter-cloud networking, infrastructure, platform, application, service management, cloud monitoring, and autonomic computing domains, respectively. This validation highlights simplicity, consistency, correctness, completeness, and usefulness of the proposed metamodel[38], [41].

7.3.4. Using Novelty Search Approach and models@runtime for Automatic Testing Environment Setup

In search-based structural testing, metaheuristic search techniques have been frequently used to automate the test data generation. In Genetic Algorithms (GAs) for example, test data are rewarded on the basis of an objective function that represents generally the number of statements or branches covered. However, owing to the wide diversity of possible test data values, it is hard to find the set of test data that can satisfy a specific coverage criterion. In this work, we introduce the use of Novelty Search (NS) algorithm to the test data generation problem based on statement-covered criteria. We believe that such approach to test data generation is attractive because it allows the exploration of the huge space of test data within the input domain. In this approach, we seek to explore the search space without regard to any objectives. In fact, instead of having a fitness-based selection, we select test cases based on a novelty score showing how different they are compared to all other solutions evaluated so far [47], [48]. We also create an architecture generation framework for setup testing environment for a distributed and heterogeneous service.

7.3.5. Using Models@Run.time to embed an Energetic Cloud Simulator in a MAPE-K Loop

Due to high electricity consumption in the Cloud datacenters, providers aim at maximizing energy efficiency through VM consolidation, accurate resource allocation or adjusting VM usage. More generally, the provider attempts to optimize resource utilization. However, while minimizing expenses, the Cloud operator still needs to conform to SLA constraints negotiated with customers (such as latency, downtime, affinity, placement, response time or duplication). Consequently, optimizing a Cloud configuration is a multi-objective problem. As a nontrivial multi-objective optimization problem, there does not exist a single solution that simultaneously optimizes each objective. There exists a (possibly infinite) number of Pareto optimal solutions. Evolutionary algorithms are popular approaches for generating Pareto optimal solutions to a multi-objective optimization problem. Most of these solutions use a fitness function to assess the quality of the candidates. However, regarding the energy consumption estimation, the fitness function can be approximative and lead to some imprecisions compared to the real observed data. This work presents a system that uses a genetic algorithm to optimize Cloud energy consumption and machine learning techniques to improve the fitness function regarding a real distributed cluster of server. We have carried out experiments on the OpenStack platform to validate our solution. This experimentation shows that the machine learning produces an accurate energy model, predicting precise values for the simulation [124][40].

7.4. Results on Diverse Implementations for Resilience

Diversity is acknowledged as a crucial element for resilience, sustainability and increased wealth in many domains such as sociology, economy and ecology. Yet, despite the large body of theoretical and experimental science that emphasizes the need to conserve high levels of diversity in complex systems, the limited amount of diversity in software-intensive systems is a major issue. This is particularly critical as these systems integrate multiple concerns, are connected to the physical world through multiple sensors, run eternally and are open to other services and to users. Here we present our latest observational and technical results about (i) new approaches to increase diversity in software systems, and (ii) software testing to assess the validity of software.

7.4.1. Software diversification

Early experiments with software diversity in the mid 1970's investigated N-version programming and recovery blocks to increase the reliability of embedded systems. Four decades later, the literature about software diversity has expanded in multiple directions: goals (fault-tolerance, security, software engineering); means (managed or automated diversity) and analytical studies (quantification of diversity and its impact). We contribute to the field of software diversity with the very first literature survey that adopts an inclusive vision of the area, with an emphasis on the most recent advances in the field. This survey includes classical work about design and data diversity for fault tolerance, as well as the cybersecurity literature that investigates randomization at different system levels. It broadens this standard scope of diversity, to include the study and exploitation of natural diversity and the management of diverse software products [17].

We also contribute to software diversity with novel techniques and methods. The interdisciplinary investigations within the DIVERSIFY project have led to the definition of novel principles for open-ended evolution in software systems. The main intuition is that software should have the ability to spontaneously and continuously evolve without waiting for specific environmental conditions. Our proposal analogizes the software consumer / provider network, which can be found in any types of distributed systems, to a bipartite ecological graph. This analogy provides the foundations for the design of an individual-based simulator used to experiment with decentralized adaptation strategies for providers and consumers. The initial model of a software network is tuned according to observations gathered from real-world software networks. The key insights about our experiments are that, 1) we can successfully model software systems as an ALife system, and 2) we succeed in emerging a global property from local decisions: when consumers and providers adapt with local decision strategies, the global robustness of the network increases. We show that these results hold with different initial situations, different scales and different topological constraints on the network [55]. In order to move towards the open-ended evolution of actual systems, we also developed a novel tool for the runtime modification of Java programs, as an extension to the JVM [60].

Our second contribution to the field of software diversity consists in experimenting its application in different fields. First, we have proposed a novel approach to exploit software diversity at multiple granularity levels simultaneously [15]. The main idea is to reconcile two aspects of the massive software reuse in web applications: on the one hand, reuse and modularity favor much writing the next killer application; on the other hand, reuse and modularity facilitates much the next massive BOBE attack. We demonstrate the feasibility of diversifying web applications at multiple levels, mitigating the risks of reuse.

The second application of automatic software diversification for Java programs aimed at answering the following question: which product line operators, applied to which program elements, can synthesize variants of programs that are incorrect, correct or perhaps even conforming to test suites? We implement source code transformations, based on the derivation operators of the Common Variability Language. We automatically synthesize more than 370,000 program variants from a set of 8 real large Java projects (up to 85,000 lines of code), obtaining an extensive panorama of the sanity of the operations [68].

The third application of software diversification is against browser fingerprinting. Browser fingerprint tracking relies on the following mechanisms: web browsers allow remote servers to discover sufficient information about a user's platform to create a digital fingerprint that uniquely identifies the platform. We argue that fingerprint uniqueness and stability are the key threats to browser fingerprint tracking, and we aim at breaking fingerprint stability over time, by exploiting software diversity and automatic reconfiguration. We leverage

virtualization and modular software architectures to automatically assemble and reconfigure a user's software components at multiple levels. We operate on the operating system, the browser, the lists of fonts and plugins. This work is the first application of software reconfiguration to build a moving target defense against browser fingerprint tracking. We have developed a prototype called *Blink* to experiment the effectiveness of our approach at randomizing fingerprints [33].

7.4.2. Software testing

Our work in the area of software testing focuses on tailoring the testing tools (analysis, generation, oracle, etc.) to specific domains. This allows us to consider domain specific knowledge (e.g., architectural patterns for GUI implementation) in order to increase the relevance and the efficiency of testing. The main results of this year are about testing GUIs and model transformations.

Graphical user interfaces (GUIs) are integral parts of software systems that require interactions from their users. Software testers have paid special attention to GUI testing in the last decade, and have devised techniques that are effective in finding several kinds of GUI errors. However, the introduction of new types of interactions in GUIs presents new kinds of errors that are not targeted by current testing techniques. We believe that to advance GUI testing, the community needs a comprehensive and high level GUI fault model, which incorporates all types of interactions. In this work, we first propose a GUI fault model designed to identify and classify GUI faults [37]. We then studied the impact of the new types of interactions in GUIs on their testing process. We show that the current GUI model-based testing approaches have limits when applied to test such new advanced GUIs [36].

Specifying a model transformation is challenging as it must be able to give a meaningful output for any input model in a possibly infinite modeling domain. Transformation preconditions constrain the input domain by rejecting input models that are not meant to be transformed by a model transformation. In our latest work [39], we present a systematic approach to discover such preconditions when it is hard for a human developer to foresee complex graphs of objects that are not meant to be transformed. The approach is based on systematically generating a finite number of test models using our tool, PRAMANA to first cover the input domain based on input domain partitioning. Tracing a transformation's execution reveals why some preconditions are missing. Using a benchmark transformation from simplified UML class diagram models to RDBMS models we discover new preconditions that were not initially specified.

We also initiated a new line of research in order to investigate Novelty Search (NS) for the automatic generation of test data. This allows the exploration of the huge space of test data within the input domain. In this approach, we select test cases based on a novelty score showing how different they are compared to all other solutions evaluated so far [47].

In Model Driven Engineering (MDE), models are first-class citizens, and model transformation is MDE's "heart and soul". Since model transformations are executed for a family of (conforming) models, their validity becomes a crucial issue. In [16] we propose to explore the question of the formal verification of model transformation properties through a tridimensional approach: the transformation involved, the properties of interest addressed, and the formal verification techniques used to establish the properties. This work is intended for a double audience. For newcomers, it provides a tutorial introduction to the field of formal verification of model transformations. For readers more familiar with formal methods and model transformations, it proposes a literature review (although not systematic) of the contributions of the field. Overall, this work allows to better understand the evolution, trends and current practice in the domain of model transformation verification. This work opens an interesting research line for building an engineering of model transformation verification guided by the notion of model transformation intent.

DYOGENE Project-Team

7. New Results

7.1. Evaluation and optimization of the quality of service perceived by mobile users for new services in cellular networks

The goal of this thesis [1] defended in 2015 is to develop tools and methods for the evaluation of the QoS (Quality of Service) perceived by users, as a function of the traffic demand, in modern wireless cellular networks. This complex problem, directly related to network dimensioning, involves modeling dynamic processes at several time-scales, which due to their randomness are amenable to probabilistic formalization. Firstly, on the ground of information theory, we capture the performance of a single link between a base station and a user in the context of a cellular network with orthogonal channels and MIMO technology. We prove and use some lower bounds of the information-theoretic ergodic capacity of such a link, which account also for the fast channel variability caused by multi-path propagation. These bounds give robust basis for further user QoS evaluation. Next, one considers several (possibly mobile) users, arriving in the network and requesting some service from it. We consider variable (elastic) bit-rate services, in which transmissions of some amounts of data are realized in a best-effort manner, or constant bit-rate services, in which a certain transmission rate needs to be maintained during requested times. On the ground of queuing theory, one captures this traffic demand and service process using appropriate (multi-class) processor sharing (PS) or loss models. In this thesis, we adapt existing PS models and develop a new loss model for wireless streaming traffic, in which the aforementioned information-theoretic capacities of single links describe the instantaneous user service rates. The multi-class models are used to capture the spatial heterogeneity of user channels, which depends on the user geographic locations and propagation shadowing phenomenon. Finally, on top of the queueing-theoretic processes, one needs to consider a multi-cellular network, whose base stations are not necessarily regularly placed, and whose geometry is further perturbed by the shadowing phenomenon. We address this randomness aspect by using some models from stochastic geometry, notably Poisson point processes and Palm formalism applied to the typical cell of the network. Applying the above three-fold approach, supposed to represent all crucial mechanisms and engineering parameters of cellular networks (such as LTE), we establish some macroscopic relations between the traffic demand and the user QoS metrics for some elastic and constant bit-rate services. These relations are mostly obtained in a semi-analytic way, i.e., they only involve static simulations of a Poisson point process (modeling the locations of base stations) in order to evaluate its characteristics which are not amenable to analytic expressions. More precisely, regarding the data traffic (the elastic bit-rate service), we capture the inter-cell interference, making the PS queue models of individual cells dependent, via some system of cell-load equations. These equations allow one to determine the mean user throughput, the mean number of users and the mean cell load in a large network, as a function of the traffic demand. The spatial distribution of these QoS metrics in the network is also studied. We validate our approach by comparing the obtained results with those measured from live-network traces. We observe a remarkably good agreement between the model predictions and the statistical data collected in several deployment scenarios. Regarding constant bit-rate services, we propose a new stochastic model to evaluate the frequency and the number of interruptions during real-time streaming calls in function of user radio conditions. Despite some fundamental similarities with the classical Erlang loss model, a more adequate model was required for in this case, where the denial of service is not definitive for a given call: it takes the form of, hopefully short, interruptions or outage periods. Our model allows one to take into account realistic implementations of the considered streaming service. We use it to study the quality of service metrics in function of user radio conditions in LTE networks. All established results contribute to the development of network dimensioning methods and are currently used in Orange internal tools for network capacity calculations.

7.2. Interference and SINR coverage in spatial non-slotted Aloha networks

In [8] we propose two analytically tractable stochastic-geometric models of interference in ad-hoc networks using pure (non-slotted) Aloha as the medium access. In contrast the slotted model, the interference in pure Aloha may vary during the transmission of a tagged packet. We develop closed form expressions for the Laplace transform of the empirical average of the interference experienced during the transmission of a typical packet. Both models assume a power-law path-loss function with arbitrarily distributed fading and feature configurations of transmitters randomly located in the Euclidean plane according to a Poisson point process. Depending on the model, these configurations vary over time or are static. We apply our analysis of the interference to study the Signal-to-Interference-and-Noise Ratio (SINR) outage probability for a typical transmission in pure Aloha. The results are used to compare the performance of non-slotted Aloha to the slotted one, which has almost exclusively been previously studied in the same context of mobile ad-hoc networks.

7.3. Random linear multihop relaying in a general field of interferers using spatial Aloha

In [9] we study, as a basic model, a stationary Poisson pattern of nodes on a line embedded in an independent planar Poisson field of interfering nodes. Assuming slotted Aloha and the signal-to-interference-and-noise ratio capture condition, with the usual power-law path loss model and Rayleigh fading, we explicitly evaluate several local and end-to-end performance characteristics related to the nearest-neighbor packet relaying on this line, and study their dependence on the model parameters (the density of relaying and interfering nodes, Aloha tuning and the external noise power). Our model can be applied in two cases: the first use is for vehicular ad-hoc networks, where vehicles are randomly located on a straight road. The second use is to study a typical route traced in a (general) planar ad-hoc network by some routing mechanism. The approach we have chosen allows us to quantify the non-efficiency of long-distance routing in pure ad-hoc networks and evaluate a possible remedy for it in the form of additional fixed relaying nodes, called road-side units in a vehicular network. It also allows us to consider a more general field of interfering nodes and study the impact of the clustering of its nodes the routing performance. As a special case of a field with more clustering than the Poisson field, we consider a Poisson-line field of interfering nodes, in which all the nodes are randomly located on random straight lines. The comparison to our basic model reveals a paradox: clustering of interfering nodes decreases the outage probability of a single (typical) transmission on the route, but increases the mean end-to-end delay.

7.4. Studying the SINR process of the typical user in Poisson networks by using its factorial moment measures

Based on a stationary Poisson point process, a wireless network model with random propagation effects (shadowing and/or fading) is considered in [7] in order to examine the process formed by the signal-to-interference-plus-noise ratio (SINR) values experienced by a typical user with respect to all base stations in the down-link channel. This SINR process is completely characterized by deriving its factorial moment measures, which involve numerically tractable, explicit integral expressions. This novel framework naturally leads to expressions for the k -coverage probability, including the case of random SINR threshold values considered in multi-tier network models. While the k -coverage probabilities correspond to the marginal distributions of the order statistics of the SINR process, a more general relation is presented connecting the factorial moment measures of the SINR process to the joint densities of these order statistics. This gives a way for calculating exact values of the coverage probabilities arising in a general scenario of signal combination and interference cancellation between base stations. The presented framework consisting of mathematical representations of SINR characteristics with respect to the factorial moment measures holds for the whole domain of SINR and is amenable to considerable model extension.

7.5. Performance laws of large heterogeneous cellular networks

In [24] we propose a model for heterogeneous cellular networks assuming a space-time Poisson process of call arrivals, independently marked by data volumes, and served by different types of base stations (having different

transmission powers) represented by the superposition of independent Poisson processes on the plane. Each station applies a processor sharing policy to serve users arriving in its vicinity, modeled by the Voronoi cell perturbed by some random signal propagation effects (shadowing). Users' peak service rates depend on their signal-to-interference-and-noise ratios (SINR) with respect to the serving station. The mutual-dependence of the cells (due to the extra-cell interference) is captured via some system of cell-load equations impacting the spatial distribution of the SINR. We use this model to study in a semi-analytic way (involving only static simulations, with the temporal evolution handled by the queuing theoretic results) network performance metrics (cell loads, mean number of users) and the quality of service perceived by the users (mean throughput) served by different types of base stations. Our goal is to identify macroscopic laws regarding these performance metrics, involving averaging both over time and the network geometry. The revealed laws are validated against real field measurement in an operational network.

7.6. Wireless networks appear Poissonian due to strong shadowing

Geographic locations of cellular base stations sometimes can be well fitted with spatial homogeneous Poisson point processes. In [6] we make a complementary observation: In the presence of the log-normal shadowing of sufficiently high variance, the statistics of the propagation loss of a single user with respect to different network stations are invariant with respect to their geographic positioning, whether regular or not, for a wide class of empirically homogeneous networks. Even in perfectly hexagonal case they appear as though they were realized in a Poisson network model, i.e., form an inhomogeneous Poisson point process on the positive half-line with a power-law density characterized by the path-loss exponent. At the same time, the conditional distances to the corresponding base stations, given their observed propagation losses, become independent and log-normally distributed, which can be seen as a decoupling between the real and model geometry. The result applies also to Suzuki (Rayleigh-log-normal) propagation model. We use Kolmogorov-Smirnov test to empirically study the quality of the Poisson approximation and use it to build a linear-regression method for the statistical estimation of the value of the path-loss exponent.

7.7. What frequency bandwidth to run cellular network in a given country? - a downlink dimensioning problem

In [25] we propose an analytic approach to the frequency bandwidth dimensioning problem, faced by cellular network operators who deploy/upgrade their networks in various geographical regions (countries) with an inhomogeneous urbanization. We present a model allowing one to capture fundamental relations between users' quality of service parameters (mean downlink throughput), traffic demand, the density of base station deployment, and the available frequency bandwidth. These relations depend on the applied cellular technology (3G or 4G impacting user peak bit-rate) and on the path-loss characteristics observed in different (urban, sub-urban and rural) areas. We observe that if the distance between base stations is kept inversely proportional to the distance coefficient of the path-loss function, then the performance of the typical cells of these different areas is similar when serving the same (per-cell) traffic demand. In this case, the frequency bandwidth dimensioning problem can be solved uniformly across the country applying the mean cell approach proposed in [Blaszczyszyn et al. WiOpt2014]. We validate our approach by comparing the analytical results to measurements in operational networks in various geographical zones of different countries.

7.8. Optimal Geographic Caching In Cellular Networks

In [23] we consider the problem of an optimal geographic placement of content in wireless cellular networks modelled by Poisson point processes. Specifically, for the typical user requesting some particular content and whose popularity follows a given law (e.g. Zipf), we calculate the probability of finding the content cached in one of the base stations. Wireless coverage follows the usual signal-to-interference-and noise ratio (SINR) model, or some variants of it. We formulate and solve the problem of an optimal randomized content placement policy, to maximize the user's hit probability. The result dictates that it is not always optimal to follow the standard policy "cache the most popular content, everywhere". In fact, our numerical results regarding three different coverage scenarios, show that the optimal policy significantly increases the chances of hit under high-coverage regime, i.e., when the probabilities of coverage by more than just one station are high enough.

7.9. Spatial distribution of the SINR in Poisson cellular networks with sector antennas

In [5] we consider a model of cellular networks where the base station locations constitute a Poisson point process and each base station is equipped with three sectorial antennas is proposed. This model permits to study the spatial distribution of the SINR in the downlink. In particular, this distribution is shown to be insensitive to the distribution of antenna azimuths. Moreover, the effect of horizontal sectorisation is shown to be equivalent to that of shadowing. Assuming ideal vertical antenna pattern, an explicit expression of the Laplace transform of the inverse of SINR is given. The model is validated by comparing its results to measurements in an operational network. It is observed numerically that, in the case of dense urban regions where interference is preponderant, one may neglect the effect of the vertical sectorization when calculating the distribution of the SINR, which provides considerable tractability. Combined with queuing theory results, the SINR's distribution permits to express the user's quality of service as function of the traffic demand. This permits in particular to operators to predict the required investments to face the continual increase of traffic demand.

7.10. Theoretical expression of link performance in OFDM cellular networks with MIMO compared to simulation and measurements

The objective of [18] is to establish a theoretical expression of the link performance in the downlink of a multiple input multiple output (MIMO) cellular network and compare it to the real Long-Term Evolution (LTE) performance. In order to account for the interference, we prove that the worst additive noise process in the MIMO context is the white Gaussian one. Based on this theoretical result, we build an analytic expression of the link performance in LTE cellular networks with MIMO. We study also the minimum mean square error (MMSE) scheme currently implemented in the field, as well as its improvement MMSE-SIC (successive interference cancellation) known to achieve the MIMO capacity. Comparison to simulation results as well as to measurements in the field shows that the theoretical expression predicts well practical link performance of LTE cellular networks. This theoretical expression of link performance is the basis of a global analytic approach to the evaluation of the quality of service perceived by the users in the long run of their arrivals and departures.

7.11. Information Theory: Boolean model in the Shannon Regime

In a paper accepted for publication in the Journal of Applied Probability, F. Baccelli and V. Anantharam consider a family of Boolean models, indexed by integers $n \geq 1$. The n -th model features a Poisson point process in \mathbb{R}^n of intensity $e^{n\rho_n}$ and balls of independent and identically distributed radii distributed like $\bar{X}_n \sqrt{n}$. Assume that $\rho_n \rightarrow \rho$ as $n \rightarrow \infty$, and that \bar{X}_n satisfies a large deviations principle. It is shown that there then exist three deterministic thresholds: τ_d the degree threshold; τ_p the percolation probability threshold; and τ_v the volume fraction threshold, such that asymptotically as n tends to infinity, we have the following features. (i) For $\rho < \tau_d$, almost every point is isolated, namely its ball intersects no other ball; (ii) for $\tau_d < \rho < \tau_p$, the mean number of balls intersected by a typical ball converges to infinity and nevertheless there is no percolation; (iii) for $\tau_p < \rho < \tau_v$, the volume fraction is 0 and nevertheless percolation occurs; (iv) for $\tau_d < \rho < \tau_v$, the mean number of balls intersected by a typical ball converges to infinity and nevertheless the volume fraction is 0; (v) for $\rho > \tau_v$, the whole space covered. The analysis of this asymptotic regime is motivated by problems in information theory, but it could be of independent interest in stochastic geometry. The relations between these three thresholds and the Shannon–Poltyrev threshold are discussed.

7.12. Stochastic Geometry: Wireless Modeling

In an Infocom'15 paper, F. Baccelli and X. Zhang (Qualcomm) have introduced an analytically tractable stochastic geometry model for urban wireless networks, where the locations of the nodes and the shadowing are highly correlated and different path loss functions can be applied to line-of-sight (LOS) and non-line-of-sight (NLOS) links.

Using a distance-based LOS path loss model and a blockage (shadowing)-based NLOS path loss model, one can derive the distribution of the interference observed at a typical location and the joint distribution at different locations of the network. When applied to cellular networks, this model leads to tractable coverage probabilities (SINR distribution) expressions. This model captures important features of urban wireless networks, which were difficult to analyze using existing models.

This model was lately extended in a joint work by the same authors and Robert Heath (UT Austin) in a paper presented at IEEE Globecom'15 where it received the best paper award.

7.13. Information Theory: SIMO

In a paper to be published in IEEE Transactions of Information Theory, F. Baccelli, N. Lee and Robert Heath consider large random wireless networks where transmit-and-receive node pairs communicate within a certain range while sharing a common spectrum. By modeling the spatial locations of nodes as Poisson point processes, analytical expressions for the ergodic spectral efficiency of a typical node pair are derived as a function of the channel state information available at a receiver (CSIR) in terms of relevant system parameters: the density of communication links, the number of receive antennas, the path loss exponent, and the operating signal-to-noise ratio. One key finding is that when the receiver only exploits CSIR for the direct link, the sum spectral efficiency increases linearly with the density, provided the number of receive antennas increases as a certain super-linear function of the density. When each receiver exploits CSIR for a set of dominant interfering links in addition to that of the direct link, the sum spectral efficiency increases linearly with both the density and the path loss exponent if the number of antennas is a linear function of the density. This observation demonstrates that having CSIR for dominant interfering links provides an order gain in the scaling law. It is also shown that this linear scaling holds for direct CSIR when incorporating the effect of the receive antenna correlation, provided that the rank of the spatial correlation matrix scales super-linearly with the density. These scaling laws are derived from integral representations of the distribution of the Signal to Interference and Noise Ratio, which are of independent interest and which in turn derived from stochastic geometry and more precisely from the theory of Shot Noise fields.

7.14. Theory of point processes

In a joint work with Mir-Omid Haji-Mirsadeghi, Sharif University, Department of Mathematics, F. Baccelli studied a class of non-measure preserving dynamical systems on counting measures called point-maps. This research introduced two objects associated with a point map f acting on a stationary point process Φ :

- The f -probabilities of Φ , which can be interpreted as the stationary regimes of the action of f on Φ . These probabilities are defined from the compactification of the action of the semigroup of point-map translations on the space of Palm probabilities. The f -probabilities of Φ are not always Palm distributions.
- The f -foliation of Φ , a partition of the support of Φ which is the discrete analogue of the stable manifold of f , i.e., the leaves of the foliation are the points of Φ with the same asymptotic fate for f . These leaves are not always stationary point processes. There always exists a point-map allowing one to navigate the leaves in a measure-preserving way.

Two papers on the matter available. The first one is under revision for Annals of Probability.

7.15. Cross-Technology Interference Mitigation in Body Area Networks: An Optimization Approach

In recent years, wearable devices and wireless body area networks have gained momentum as a means to monitor people's behavior and simplify their interaction with the surrounding environment, thus representing a key element of the body-to-body networking (BBN) paradigm. Within this paradigm, several transmission technologies, such as 802.11 and 802.15.4, that share the same unlicensed band (namely, the industrial, scientific, and medical band) coexist, dramatically increasing the level of interference and, in turn, negatively

affecting network performance. In this paper, we analyze the cross-technology interference (CTI) caused by the utilization of different transmission technologies that share the same radio spectrum. We formulate an optimization model that considers internal interference, as well as CTI to mitigate the overall level of interference within the system, explicitly taking into account node mobility. We further develop three heuristic approaches to efficiently solve the interference mitigation problem in large-scale network scenarios. Finally, we propose a protocol to compute the solution that minimizes CTI in a distributed fashion. Numerical results show that the proposed heuristics represent efficient and practical alternatives to the optimal solution for solving the CTI mitigation (CTIM) problem in large-scale BBN scenarios.

7.16. Body-to-Body Area Networks

The ongoing evolution of wireless technologies has fostered the development of innovative network paradigms like the Internet of Things (IoT). Wireless Body Area Networks, and more specifically Body-to-Body Area Networks (BBNs), are emerging solutions for the monitoring of people's behavior and their interaction with the surrounding environment. These networks represent a key building block of the IoT paradigm. In BBNs several transmission technologies like 802.11 and 802.15.4 that share the same unlicensed band (namely the industrial, scientific and medical (ISM) radio band) coexist, increasing dramatically the level of interference and, in turn, negatively affecting network's performance. In [14], we investigate the Cross-Technology Interference Mitigation (CTIM) problem caused by the utilization of different transmission technologies that share the same radio spectrum, from a centralized and distributed point of view, respectively.

7.17. Exact Worst-Case Delay in FIFO-Multiplexing Feed-Forward Networks

In [11], we compute the actual worst-case end-to-end delay for a flow in a feed-forward network of FIFO-multiplexing service curve nodes, where flows are shaped by piecewise-affine concave arrival curves, and service curves are piecewise affine and convex. We show that the worst-case delay problem can be formulated as a mixed integer-linear programming problem, whose size grows exponentially with the number of nodes involved. Furthermore, we present approximate solution schemes to find upper and lower delay bounds on the worst-case delay. Both only require to solve just one linear programming problem, and yield bounds which are generally more accurate than those found in the previous work, which are computed under more restrictive assumptions.

7.18. Fast symbolic computation of the worst-case delay in tandem networks and applications

Computing deterministic performance guarantees is a defining issue for systems with hard real-time constraints, like reactive embedded systems. In [10], we use burst-rate constrained arrivals and rate-latency servers to deduce tight worst-case delay bounds in tandem networks under arbitrary multiplexing. We present a constructive method for computing the exact worst-case delay, which we prove to be a linear function of the burstiness and latencies; our bounds are hence symbolic in these parameters. Our algorithm runs in quadratic time in the number of servers. We also present an application of our algorithm to the case of stochastic arrivals and server capacities. For a generalization of the exponentially bounded burstiness (EBB) model, we deduce a polynomial-time algorithm for stochastic delay bounds that strictly improve the state-of-the-art separated flow analysis (SFA) type bounds.

7.19. Ancillary Service to the Grid Using Intelligent Deferrable Loads

Renewable energy sources such as wind and solar power have a high degree of unpredictability and time-variation, which makes balancing demand and supply challenging. One possible way to address this challenge is to harness the inherent flexibility in demand of many types of loads. Introduced in [19] is a technique for decentralized control for automated demand response that can be used by grid operators as ancillary service for maintaining demand-supply balance. A randomized control architecture is proposed, motivated by the need for decentralized decision making, and the need to avoid synchronization that can lead to large and detrimental

spikes in demand. An aggregate model for a large number of loads is then developed by examining the mean field limit. A key innovation is a linear time-invariant (LTI) system approximation of the aggregate nonlinear model, with a scalar signal as the input and a measure of the aggregate demand as the output. This makes the approximation particularly convenient for control design at the grid level.

7.20. Spectral Decomposition of Demand-Side Flexibility for Reliable Ancillary Services in a Smart Grid

[22] describes a new way of thinking about demand-side resources to provide ancillary services to control the grid. It is shown that loads can be classified based on the frequency bandwidth of ancillary service that they can offer. If demand response from loads respects these frequency limitations, it is possible to obtain highly reliable ancillary service to the grid, while maintaining strict bounds on the quality of service (QoS) delivered by each load. It is argued that automated demand response is required for reliable control. Moreover, some intelligence is needed at demand response loads so that the aggregate will be reliable and controllable.

7.21. State Estimation for the Individual and the Population in Mean Field Control with Application to Demand Dispatch

[29] concerns state estimation problems in a mean field control setting. In a finite population model, the goal is to estimate the joint distribution of the population state and the state of a typical individual. The observation equations are a noisy measurement of the population. The general results are applied to demand dispatch for regulation of the power grid, based on randomized local control algorithms. In prior work by the authors it has been shown that local control can be carefully designed so that the aggregate of loads behaves as a controllable resource with accuracy matching or exceeding traditional sources of frequency regulation. The operational cost is nearly zero in many cases. The information exchange between grid and load is minimal, but it is assumed in the overall control architecture that the aggregate power consumption of loads is available to the grid operator. It is shown that the Kalman filter can be constructed to reduce these communication requirements, and to provide the grid operator with accurate estimates of the mean and variance of quality of service (QoS) for an individual load.

7.22. Perfect sampling of Jackson queueing networks

In [12], we consider open Jackson networks with losses with mixed finite and infinite queues and analyze the efficiency of sampling from their exact stationary distribution. We show that perfect sampling is possible, although the underlying Markov chain may have an infinite state space. The main idea is to use a Jackson network with infinite buffers (that has a product form stationary distribution) to bound the number of initial conditions to be considered in the coupling from the past scheme. We also provide bounds on the sampling time of this new perfect sampling algorithm for acyclic or hyper-stable networks. These bounds show that the new algorithm is considerably more efficient than existing perfect samplers even in the case where all queues are finite. We illustrate this efficiency through numerical experiments. We also extend our approach to variable service times and non-monotone networks such as queueing networks with negative customers.

7.23. Speeding up Glauber Dynamics for Random Generation of Independent Sets

The maximum independent set (MIS) problem is a well-studied combinatorial optimization problem that naturally arises in many applications, such as wireless communication, information theory and statistical mechanics. MIS problem is NP-hard, thus many results in the literature focus on fast generation of maximal independent sets of high cardinality. One possibility is to combine Gibbs sampling with coupling from the past arguments to detect convergence to the stationary regime. This results in a sampling procedure with time complexity that depends on the mixing time of the Glauber dynamics Markov chain. We propose in [37] an adaptive method for random event generation in the Glauber dynamics that considers only the events that are effective in the coupling from the past scheme, accelerating the convergence time of the Gibbs sampling algorithm.

7.24. Approximate optimality with bounded regret in dynamic matching models

In [28], we consider a dynamic matching model with random arrivals. In prior work, authors have proposed policies that are stabilizing, and also policies that are approximately finite-horizon optimal. This paper considers the infinite-horizon average-cost optimal control problem. A relaxation of the stochastic control problem is proposed, which is found to be a special case of an inventory model, as treated in the classical theory of Clark and Scarf. The optimal policy for the relaxation admits a closed-form expression. Based on the policy for this relaxation, a new matching policy is proposed. For a parameterized family of models in which the network load approaches capacity, this policy is shown to be approximately optimal, with bounded regret, even though the average cost grows without bound.

7.25. Perfect sampling for multiclass closed queueing networks

In [27] we present an exact sampling method for multiclass closed queueing networks. We consider networks for which stationary distribution does not necessarily have a product form. The proposed method uses a compact representation of sets of states, that is used to derive a bounding chain with significantly lower complexity of one-step transition in the coupling from the past scheme. The coupling time of this bounding chain can be larger than the coupling time of the exact chain, but it is finite in expectation. Numerical experiments show that coupling time is close to that of the exact chain. Moreover, the running time of the proposed algorithm outperforms the classical algorithm.

7.26. Fast and Memory Optimal Low-Rank Matrix Approximation

In this paper, we revisit the problem of constructing a near-optimal rank k approximation of a matrix $M \in [0, 1]^{m \times n}$ under the streaming data model where the columns of M are revealed sequentially. We present SLA (Streaming Low-rank Approximation), an algorithm that is asymptotically accurate, when $ks_{k+1}(M) = o(\sqrt{mn})$ where $s_{k+1}(M)$ is the $(k+1)$ -th largest singular value of M . This means that its average mean-square error converges to 0 as m and n grow large (i.e., $\|\widehat{M}^{(k)} - M^{(k)}\|_F^2 = o(mn)$ with high probability, where $\widehat{M}^{(k)}$ and $M^{(k)}$ denote the output of SLA and the optimal rank k approximation of M , respectively). Our algorithm makes one pass on the data if the columns of M are revealed in a random order, and two passes if the columns of M arrive in an arbitrary order. To reduce its memory footprint and complexity, SLA uses random sparsification, and samples each entry of M with a small probability δ . In turn, SLA is memory optimal as its required memory space scales as $k(m+n)$, the dimension of its output. Furthermore, SLA is computationally efficient as it runs in $O(\delta kmn)$ time (a constant number of operations is made for each observed entry of M), which can be as small as $O(k \log(m)^4 n)$ for an appropriate choice of δ and if $n \geq m$.

7.27. Combinatorial Bandits Revisited

[42] investigates stochastic and adversarial combinatorial multi-armed bandit problems. In the stochastic setting under semi-bandit feedback, we derive a problem-specific regret lower bound, and discuss its scaling with the dimension of the decision space. We propose ESCB, an algorithm that efficiently exploits the structure of the problem and provide a finite-time analysis of its regret. ESCB has better performance guarantees than existing algorithms, and significantly outperforms these algorithms in practice. In the adversarial setting under bandit feedback, we propose COMBEXP, an algorithm with the same regret scaling as state-of-the-art algorithms, but with lower computational complexity for some combinatorial problems.

7.28. Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs

A non-backtracking walk on a graph is a directed path such that no edge is the inverse of its preceding edge. The non-backtracking matrix of a graph is indexed by its directed edges and can be used to count non-backtracking walks of a given length. It has been used recently in the context of community detection and has appeared previously in connection with the Ihara zeta function and in some generalizations of Ramanujan graphs. In [26], we study the largest eigenvalues of the non-backtracking matrix of the Erdos-Renyi random graph and of the Stochastic Block Model in the regime where the number of edges is proportional to the number of vertices. Our results confirm the "spectral redemption" conjecture that community detection can be made on the basis of the leading eigenvectors above the feasibility threshold.

7.29. Designing Adaptive Replication Schemes in Distributed Content Delivery Networks

In [32], we address the problem of content replication in large distributed content delivery networks, composed of a data center assisted by many small servers with limited capabilities and located at the edge of the network. The objective is to optimize the placement of contents on the servers to offload as much as possible the data center. We model the system constituted by the small servers as a loss network, each loss corresponding to a request to the data center. Based on large system / storage behavior, we obtain an asymptotic formula for the optimal replication of contents and propose adaptive schemes related to those encountered in cache networks but reacting here to loss events, and faster algorithms generating virtual events at higher rate while keeping the same target replication. We show through simulations that our adaptive schemes outperform significantly standard replication strategies both in terms of loss rates and adaptation speed.

7.30. Spectral Detection in the Censored Block Model

In [36], we consider the problem of partially recovering hidden binary variables from the observation of (few) censored edge weights, a problem with applications in community detection, correlation clustering and synchronization. We describe two spectral algorithms for this task based on the non-backtracking and the Bethe Hessian operators. These algorithms are shown to be asymptotically optimal for the partial recovery problem, in that they detect the hidden assignment as soon as it is information theoretically possible to do so.

7.31. A spectral method for community detection in moderately-sparse degree-corrected stochastic block models

In the ordinary stochastic block model, all degrees in a cluster have the same expected degree. The Degree-Corrected Stochastic Block Models (DC-SBM) is a generalization of the former where the expected degrees of individual nodes follow a prescribed degree-sequence. We consider community detection in the DC-SBM in a paper currently in preparation [43]. We perform spectral clustering on a suitably normalized adjacency matrix. This leads to consistent recovery of the block-membership of all but a vanishing fraction of nodes, in the regime where the lowest degree is of order $\log(n)$ or higher. The main contributions of this paper are (i) the fact that recovery succeeds for very heterogeneous degree-distributions and (ii) a clean analysis for the DC-SBM, which is a messy model.

7.32. An Impossibility Result for Reconstruction in a Degree-Corrected Planted-Partition Model

In a paper currently in preparation [44], we consider a degree-corrected planted-partition model: a random graph on n nodes with two equal-sized clusters. The model parameters are two constants $a, b > 0$ and an i.i.d. sequence $(\phi_i)_{i=1}^n$, with finite second moment Φ^2 . Vertices i and j are joined by an edge with probability $\frac{\phi_i \phi_j}{n} a$ whenever they are in the same class and with probability $\frac{\phi_i \phi_j}{n} b$ otherwise. We prove that the underlying community structure cannot be accurately recovered from observations of the graph when $(a - b)^2 \Phi^2 \leq 2(a + b)$.

7.33. Universality in polytope phase transitions and message passing algorithms

In [], we consider a class of nonlinear mappings $F_{A,N}$ in \mathbb{R}^N indexed by symmetric random matrices $A \in \mathbb{R}^{N \times N}$ with independent entries. Within spin glass theory, special cases of these mappings correspond to iterating the TAP equations and were studied by Bolthausen [Comm. Math. Phys. 325 (2014) 333-366]. Within information theory, they are known as "approximate message passing" algorithms. We study the high-dimensional (large N) behavior of the iterates of F for polynomial functions F , and prove that it is universal; that is, it depends only on the first two moments of the entries of A , under a sub-Gaussian tail condition. As an application, we prove the universality of a certain phase transition arising in polytope geometry and compressed sensing. This solves, for a broad class of random projections, a conjecture by David Donoho and Jared Tanner.

7.34. Contagions in Random Networks with Overlapping Communities

In [13], we consider a threshold epidemic model on a clustered random graph with overlapping communities. In other words, our epidemic model is such that an individual becomes infected as soon as the proportion of her infected neighbors exceeds the threshold q of the epidemic. In our random graph model, each individual can belong to several communities. The distributions for the community sizes and the number of communities an individual belongs to are arbitrary. We consider the case where the epidemic starts from a single individual, and we prove a phase transition (when the parameter q of the model varies) for the appearance of a cascade, i.e. when the epidemic can be propagated to an infinite part of the population. More precisely, we show that our epidemic is entirely described by a multi-type (and alternating) branching process, and then we apply Sevastyanov's theorem about the phase transition of multi-type Galton-Watson branching processes. In addition, we compute the entries of the matrix whose largest eigenvalue gives the phase transition.

7.35. The Diameter of Weighted Random Graphs.

In [3], we study the impact of random exponential edge weights on the distances in a random graph and, in particular, on its diameter. Our main result consists of a precise asymptotic expression for the maximal weight of the shortest weight paths between all vertices (the weighted diameter) of sparse random graphs, when the edge weights are i.i.d. exponential random variables.

EVA Team

7. New Results

7.1. Wireless Sensor Networks

7.1.1. Time slot and channel assignment in multichannel Wireless Sensor Networks

Participants: Pascale Minet, Ridha Soua, Erwan Livolant.

Wireless sensor networks (WSNs) play a major role in industrial environments for data gathering (convergecast). Among the industrial requirements, we can name a few like 1) determinism and bounded convergecast latencies, 2) throughput and 3) robustness against interferences. The classical IEEE 802.15.4 that has been designed for low power lossy networks (LLNs) partially meets these requirements. That is why the IEEE 802.15.4e MAC amendment has been proposed recently. This amendment combines a slotted medium access with a channel hopping (i.e. Time Slotted Channel Hopping TSCH). The MAC layer orchestrates the medium accesses of nodes according to a given schedule. Nevertheless, this amendment does not specify how this schedule is computed. We propose a distributed joint time slot and channel assignment, called *Wave* for data gathering in LLNs. This schedule targets minimized data convergecast delays by reducing the number of slots assigned to nodes. Moreover, *Wave* ensures the absence of conflicting transmissions in the schedule provided. In such a schedule, a node is awake only during its slots and the slots of its children in the convergecast routing graph. Thus, energy efficiency is ensured. We describe in details the functioning of *Wave*, highlighting its features (e.g. support of heterogeneous traffic, support of a sink equipped with multiple interfaces) and properties in terms of worst case delays and buffer size. We discuss its features with regard to a centralized scheduling algorithm like *TMCP* and a distributed one like *DeTAS*. Simulation results show the good performance of *Wave* compared to *TMCP*. Since in an industrial environment, several routing graphs can coexist, we study how *Wave* supports this coexistence.

7.1.2. Centralized Scheduling in TSCH-based Wireless Sensor Networks

Participants: Erwan Livolant, Pascale Minet, Thomas Watteyne.

Scheduling in an IEEE802.15.4e TSCH(Time Slotted Channel Hopping 6TiSCH) low-power wireless network can be done in a centralized or distributed way. When using centralized scheduling, a scheduler installs a communication schedule into the network. This can be done in a standards-based way using CoAP. In this study, we compute the number of packets and the latency this takes, on real-world examples. The result is that the cost is very high using today's standards, much higher than when using an ad-hoc solution such as OCARI. We conclude by making recommendations to drastically reduce the number of messages and improve the efficiency of the standardized approach.

7.1.3. Distributed and Optimized Deployment of WSNs

Participants: Ines Khoufi, Pascale Minet.

This is a joint work with Telecom SudParis: Anis Laouiti.

We are witnessing the deployment of many wireless sensor networks in various application domains such as pollution detection in the environment, intruder detection at home, preventive maintenance in industrial process, monitoring of temporary industrial worksites, damage assessment after a disaster.... Wireless sensor networks are deployed to monitor physical phenomena. The accuracy of the information collected depends on the position of sensor nodes. These positions must meet the application requirements in terms of coverage and connectivity, which therefore requires the use of deployment algorithms. We distinguish two cases: firstly when the nodes are autonomous, and secondly when they are static and the deployment is assisted by mobile robots. In both cases, this deployment must not only meet the application's coverage and connectivity requirements, but must also minimize the number of sensors needed while satisfying various constraints (e.g. obstacles, energy, fault-tolerant connectivity). We distinguished two cases: autonomous and mobile wireless sensor nodes on the one hand, and static wireless sensor nodes on the other hand.

We propose a distributed and optimized deployment of mobile and autonomous sensor nodes to ensure full coverage of the 2D-area considered, as well as network connectivity. With the full coverage of the area, any event occurring in this area is detected by at least one sensor node. In addition, the connectivity ensures that this event is reported to the sink in charge of analyzing the data gathered from the sensors and acting according to these data. This distributed algorithm, called OA-DVFA, can run in an unknown area with obstacles discovered dynamically. We distinguish two types of obstacles: the transparent ones like ponds in outdoor environment, or tables in an indoor site that only prevent the location of sensor nodes inside them; whereas the opaque obstacles like walls or trees prevent the sensing by causing the existence of hidden zones behind them: such zones may remain uncovered. Opaque obstacles are much more complex to handle than transparent ones and require the deployment of additional sensors to eliminate coverage holes. OA-DVFA is based on virtual forces to obtain a fast spreading of sensor nodes and uses a virtual grid to stop node oscillations and save energy by making sleep redundant nodes. It automatically detects when the maximum area coverage is reached.

We also considered 3D volumes and proposed an algorithm, called 3D-DVFA, also based on virtual forces, to ensure full coverage of 3D volumes and ensure network connectivity. This is a joint work with Nadya Boufares from ENSI, Tunisia. Since applications of such 3D deployments may be limited, we focus on 3D surface covering, where the objective is to deploy wireless sensor nodes on a 3D-surface (e.g. a mountain) to ensure full area coverage and network connectivity. To reach this goal we propose 3D-DVFA-SC, a distributed deployment algorithm based on virtual forces strategy to move sensor nodes.

7.1.4. WSN deployment assisted by mobile robots

Participants: Ines Khoufi, Pascale Minet.

This is a joint work with Telecom SudParis: Anis Laouiti.

Autonomous deployment may be expensive when the number of mobile sensor nodes is very high. In this case, an assisted deployment may be necessary: the nodes' positions being pre-computed and given to mobile robots that place a static sensor at each position. In order to reduce both the energy consumed by the robots, their exposure time to a hostile environment, as well as the time at which the wireless network becomes operational, the optimal tour of robots is this minimizing the delay. This delay must take into account not only the time needed by the robot to travel the tour distance but also the time spent in the rotations performed by the robot each time it changes its direction. This problem is called the Multiple Robot Deploying Sensor nodes problem, in short MRDS. We first show how this problem differs from the well-known traveling salesman problem. We adopt two approaches to optimize the deployment duration. The first one is based on game theory to optimize the length of the tours of two robots (TRDS), and the second is based on a multi-objective optimization, for multiple robots (MRDS). The objectives to be met are: optimizing the duration of the longest tour, balancing the durations of the robot tours and minimizing the number of robots used, while bypassing obstacles.

The TRDS problem is modeled as a non-cooperative game with two players representing the mobile robots, these robots compete for the selection of the sensor nodes to deploy. Each robots tends to maximize its utility function.

We then propose an integer linear program formulation of the MRDS problem. We propose various algorithms relevant to iterative improvement by exchanging tour edges, genetic approach and hybridization. The solutions provided by these algorithms are compared and their closeness to the optimal is evaluated in various configurations.

7.1.5. Sinks Deployment and Packet Scheduling for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

The objective of this work is to propose an optimal deployment and distributed packet scheduling of multi-sink Wireless Sensors networks (WNSs). We start by computing the optimal deployment of sinks for a given maximum number of hops between nodes and sinks. We also propose an optimal distributed packet scheduling in order to estimate the minimum energy consumption. We consider the energy consumed due to reporting, forwarding and overhearing. In contrast to reporting and forwarding, the energy used in overhearing is difficult to estimate because it is dependent on the packet scheduling. In this case, we determine the lower-bound of overhearing, based on an optimal distributed packet scheduling formulation. We also propose another estimation of the lower-bound in order to simulate non interfering parallel transmissions which is more tractable in large networks. We note that overhearing largely predominates in energy consumption. A large part of the optimizations and computations carried out in this work are obtained using ILP (Integer Linear Programming) formalization.

7.1.6. Security in wireless sensor networks

Participants: Selma Boumerdassi, Paul Muhlethaler.

Sensor networks are often used to collect data from the environment where they are located. These data can then be transmitted regularly to a special node called a *sink*, which can be fixed or mobile. For critical data (like military or medical data), it is important that sinks and simple sensors can mutually authenticate so as to avoid data to be collected and/or accessed by fake nodes. For some applications, the collection frequency can be very high. As a result, the authentication mechanism used between a node and a sink must be fast and efficient both in terms of calculation time and energy consumption. This is especially important for nodes which computing capabilities and battery lifetime are very low. Moreover, an extra effort has been done to develop alternative solutions to secure, authenticate, and ensure the confidentiality of sensors, and the distribution of keys in the sensor network. Specific researches have also been conducted for large-scale sensors. At present, we work on an exchange protocol between sensors and sinks based on low-cost shifts and xor operations.

7.1.7. Massive MIMO Cooperative Communications for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

This work is a collaboration with Mérouane Debbah (Supelec, France).

The objective of this work is to propose a framework for massive MIMO cooperative communications for Wireless Sensor Networks. Our main objective is to analyze the performances of the deployment of a large number of sensors. This deployment should cope with a high demand for real time monitoring and should also take into account energy consumption. We have assumed a communication protocol with two phases: an initial training period followed by a second transmit period. The first period allows the sensors to estimate the channel state and the objective of the second period is to transmit the data sensed. We start analyzing the impact of the time devoted to each period. We study the throughput obtained with respect to the number of sensors when there is one sink. We also compute the optimal number of sinks with respect to the energy spent for different values of sensors. This work is a first step to establish a complete framework to study energy efficient Wireless Sensor Networks where the sensors collaborate to send information to a sink. Currently, we are exploring the multi-hop case.

7.2. Cognitive Radio Networks

7.2.1. Multichannel time slot assignment in Cognitive Radio Sensor Networks

Participants: Ons Mabrouk, Pascale Minet.

This is a joint work with Hanen Idoudi and Leila Saidane from ENSI, Tunisia.

The unlicensed spectrum bands become overcrowded causing an increased level of interference for current wireless sensor nodes. Cognitive Radio Sensor Networks (CRSNs) overcome this problem by allowing sensor nodes to access opportunistically the underutilized licensed spectrum bands. The sink assigns the spectrum holes to the secondary users (SUs). Therefore, it must rely on reliable information about the spectrum holes to protect the primary users (PUs). We focused on the MultiChannel Time Slot Assignment problem in CRSN and tackled this problem: first at the level of a cluster (i.e. Intra-cluster multichannel scheduling), second at the level of several clusters coexisting in the same CRSN (i.e. inter-cluster multichannel scheduling).

In 2013, we proposed an Opportunistic centralized Time slot assignment in COgnitive Radio sensor networks (OTICOR) for the Intra-cluster multichannel scheduling. OTICOR differs from the existing schemes in its ability to allow non-interfering cognitive sensors to access the same channel and time slot pair. OTICOR takes advantages of spatial reuse, multichannel communication and multiple radio interfaces of the sink. We proved through simulations that a smaller schedule length improves the throughput. Applying OTICOR, we showed that, even in the presence of several *PU*s, the average throughput granted to *SU*s remains important. We also showed how to get the best performances of OTICOR when the channel occupancy by *PU*s is known.

In 2014, we extended this Intra-cluster multichannel scheduling algorithm by proposing two ways for the sink to determine the available channels and alert the *SU*s if an unexpected activity of *PU* occurs. Our objective is to design an algorithm able to detect the unexpected presence of *PU*s in the multi-hop network while maximizing the throughput. If the estimation of *PU* presence is accurate, a channel sensing at the beginning of the slotframe is sufficient. This algorithm, called Frame-ICMS (Frame Intra-Cluster Multichannel Scheduling), takes advantage of the slots dedicated to the control period by allowing noninterfering cognitive sensors to access the control/data channel and time slot pair. We showed through simulations that using the control period also for data transmission minimizes the schedule length and maximizes the throughput. However, if the estimation of *PU* presence is not accurate, channel sensing should be done before each slot. We proposed the Slot-ICMS algorithm.

In 2015, we focused on inter-cluster multichannel scheduling algorithm. We considered the coexistence of different clusters in a same CSRN, each cluster having an intra-cluster multichannel scheduling algorithm. Our goal is to obtain a better scalability without losing the properties provided by OTICOR:

- collision-free schedule,
- minimized data gathering delays,
- sleeping periods per node to save node's energy.

However, the co-existence of several clusters in the same environment may lead to conflicts in the allocation of time slots and channels between these clusters. To avoid inter-cluster collisions, we do not require that different clusters use different channels, we assign distinct channels only to nodes having one-hop neighbors out of their cluster. Once the problem of inter-cluster collision is avoided, each cluster head schedules the transmissions of its members independently. This whole solution exhibits good performances as shown by the simulation results.

7.3. Learning for an efficient and dynamic management of network resources and services

7.3.1. Learning in networks

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet, Selma Boumerdassi.

This work is a joint work with Dominique Barth (University of Versailles-Saint-Quentin). To guarantee an efficient and dynamic management of network resources and services we intend to use a powerful mathematical tool: prediction and learning from prediction. Prediction will be concerned with guessing the evolution of network or network components state, based on knowledge about the past elements and/or other available information. Basically, the prediction problem could be formulated as follows: a forecaster observes the values of one or several metrics giving indications about the network state (generally speaking the network represents the environment). At each time t , before the environment reveals the new metric values, the forecaster predicts the new values based on previous observations. Contrary to classical methods where the environment evolution is characterized by stochastic process, we suppose that the environment evolution follows an unspecified mechanism, which could be deterministic, stochastic, or even adaptive to a given behavior. The prediction process should adapt to unpredictable network state changes due to its non-stationary nature. To properly address the adaptivity challenge, a special type of forecasters is used: the experts. These experts analyse the previous environment values, apply their own computation and make their own prediction. The experts predictions are given to the forecaster before the next environment values are revealed.

The forecaster can then make its own prediction depending on the experts' "advice". The risk of a prediction may be defined as the value of a loss function measuring the discrepancy between the predicted value and the real environment value. The principal notion to optimize the behavior of the forecasters is the regret, seen as a difference between the forecaster's accumulated loss and that of each expert. To optimize the prediction process means to construct a forecasting strategy that guarantees a small loss with respect to defined experts. Adaptability of the forecaster is reflected in the manner in which it is able to follow the better expert according to the context.

Our purpose is to apply on-line learning strategies to:

- Wireless Sensor Networks (WSNs) to predict the quality of a wireless link in a WSN, based on the LQI metric for instance and take advantage of wireless links with the best possible quality to improve the packet delivery rate. We model this problem as a forecaster prediction game based on the advice of several experts. The forecaster learns on-line how to adjust its prediction to better fit the environment metric values. A forecaster estimates the LQI value using the advice of experts.
- Content Delivery Networks (CDNs) to predict the number of solicitations of video contents to cache the contents with the highest popularity.
- Data centers require a huge amount of energy. As an example, in 2014, the electric consumption of all data centers will be larger than 42 TWh, and after 2020 the CO₂ production will be larger than 1.27 Gtons, ie. more than the aeronautic industry (GeSI SMARTer 2020 report). These "frightening" figures led the research community to work on the management of energy consumption. Several tracks have been explored, among which the optimization of computation and load balancing of servers. At present, we work on tools dedicated to traffic prediction, thus allowing a better management of servers. Our work consists in modeling the traffic specific to data centers and apply different statistical prediction methods.

7.3.2. Tools for learning and prediction

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet.

In 2015, Nesrine Ben Hassine developed an extraction tool to provide real traces from YouTube. these real traces are used as a learning sample by the different prediction algorithms used.

Nesrine Ben Hassine and Dana Marinca extended their simulation tool developed in Python to integrate:

- various prediction strategies SES (Single Exponential Smoothing), DES (Double Exponential Smoothing), Basic and enhanced basic, strategies based on averages (e.g. Average on a Moving Window), regressions (e.g. polynomial or Savitzky Golay), as well as prediction strategies adapting dynamically their parameters according to the loss obtained.
- various loss functions (e.g. absolute value, square). The prediction accuracy is evaluated by a loss function as the discrepancy between the prediction value and the real number obtained.
- different forecaster strategies: Best expert, Exponential Weighted Average, K Best-Experts, etc.

With these tools, we can now tune parameters of prediction strategies and evaluate them.

7.3.3. Popularity prediction in CDNs

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet.

To predict the popularity of video contents, expressed as the number of solicitations, we compared three prediction strategies: Single Exponential Smoothing (SES), Double Exponential Smoothing (DES) and Basic. The best tuning of each strategy is determined, depending on the considered phase of the solicitation curve. For DES, values of the smoothing factor close to 1 provide the best results. We study the behavior of each strategy within a phase and around a phase change, where a phase is defined as an interval of time during which a measured metric remains relatively stable.

Basic expert makes large errors at the phase change, but it quickly corrects its prediction and it is the expert having the closest prediction to the real value within a phase. DES expert provides also good quality predictions within a phase. Since DES and Basic experts outperform the SES expert, we recommend the use of on the one hand, the best DES expert per phase within a phase and on the other hand, the Basic expert to automatically detect phase changes, because of its better reactivity. This self-learning and prediction method can be applied to optimize resources allocation in service oriented architectures and self-adaptive networks, more precisely for cache management in CDNs.

7.3.4. Automatic phase detection in popularity evolution of video contents

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet.

In Content Delivery Networks (CDNs) where experts predict the number of solicitations of video contents, simulations based on real YouTube traces show that the accuracy of prediction is improved by splitting the video content profile in contiguous phases. A phase is an interval of time during which a measured metric remains relatively stable. The best expert per phase outperforms the best expert on the whole video content profile. Different prediction methods are compared and also different phase change-points detection methods are evaluated:

- the R tool using Bayesian inference,
- the Basic expert (an important loss may indicate a phase change),
- a fixed time interval (e.g. each week).

The goal is to identify the method (or method parameters) minimizing the cumulated discrepancy compared to real solicitations of video contents. The use of this machine learning method allows the Content Delivery Network to self-adapt to users solicitations by caching the most popular contents near the end users. More generally, such method can be applied to decide which contents should be replicated to improve the performance of audio and video applications and maximize the satisfaction degree of users.

7.4. VANETs

7.4.1. Protocols for VANETs

Participants: Nadjib Achir, Younes Bouchaala, Mohamed Elhadad Or Hadded, Paul Muhlethaler, Oyunchimeg Shagdar.

7.4.1.1. Synthetic study of TDMA protocols for VANETs

Recently several Time Division Multiple Access (TDMA)-based medium access control protocols have been proposed for VANETs in an attempt to ensure that all the vehicles have enough time to send safety messages without collisions and to reduce the end-to-end delay and the packet loss ratio. In this paper, we identify the reasons for using the collision-free medium access control paradigm in VANETs. We then present a novel topology-based classification and we provide an overview of TDMA-based MAC protocols that have been proposed for VANETs. We focus on the characteristics of these protocols, as well as on their benefits and limitations. Finally, we give a qualitative comparison, and we discuss some open issues that need to be tackled in future studies in order to improve the performance of TDMA-based MAC protocols for vehicle to vehicle (V2V) communications.

7.4.1.2. A stable clustering protocol for VANETs

VANETs have a highly dynamic and portioned network topology due to the constant and rapid movement of vehicles. Currently, clustering algorithms are widely used as the control schemes to make VANET topology less dynamic for Medium Access Control (MAC), routing and security protocols. An efficient clustering algorithm must take into account all the necessary information related to node mobility. In this paper, we propose an Adaptive Weighted Clustering Protocol (AWCP), specially designed for vehicular networks, which takes the highway ID, direction of vehicles, position, speed and the number of neighboring vehicles into account in order to enhance the stability of the network topology. However, the multiple control parameters of our AWCP, make parameter tuning a nontrivial problem. In order to optimize the protocol, we define a

multi-objective problem whose inputs are the AWCP's parameters and whose objectives are: providing stable cluster structures, maximizing data delivery rate, and reducing the clustering overhead. We address this multi-objective problem with the Nondominated Sorted Genetic Algorithm version 2 (NSGA-II). We evaluate and compare its performance with other multi-objective optimization techniques: Multi-objective Particle Swarm Optimization (MOPSO) and Multi-objective Differential Evolution (MODE). The experiments reveal that NSGA-II improves the results of MOPSO and MODE in terms of spacing, spread, ratio of non-dominated solutions, and inverse generational distance, which are the performance metrics used for comparison.

7.4.1.3. Using Road IDs to Enhance Clustering in Vehicular Ad hoc Networks

Vehicular ad hoc networks (VANETs) where vehicles act as mobile nodes is an instance of Mobile Ad hoc NETWORKS (MANETs), which are essentially developed for intelligent transportation systems. A challenging problem when designing communication protocols in VANETs is coping with high vehicle mobility, which causes frequent changes in the network topology and leads to frequent breaks in communication. The clustering technique is being developed to reduce the impact of mobility between neighboring vehicles. In this paper, we propose an Adaptive Weighted Cluster Protocol for VANETs, which is a road map dependent and uses road IDs and movement direction in order to make the clusters structure as stable as possible. The experimental results reveal that AWCP outperforms four other most commonly used clustering protocols in terms of control packet overhead, the packet delivery ratio, and the average cluster lifetime, which are the most usual metrics used for comparing performance.

7.4.2. Models for VANETs

Participants: Nadjib Achir, Younes Bouchaala, Guy Fayolle, Paul Muhlethaler, Oyunchimeg Shagdar.

7.4.2.1. Model of IEEE 802.11 broadcast scheme with infinite queue

We have analyzed the so-called back-off technique of the IEEE 802.11 protocol in broadcast mode with waiting queues. In contrast to existing models, packets arriving when a station (or node) is in back-off state are not discarded, but are stored in a buffer of infinite capacity. As in previous studies, the key point of our analysis hinges on the assumption that the time on the channel is viewed as a random succession of transmission slots (whose duration corresponds to the length of a packet) and mini-slots during which the back-off of the station is decremented. These events occur independently, with given probabilities. The state of a node is represented by a two-dimensional Markov chain in discrete-time, formed by the back-off counter and the number of packets at the station. Two models are proposed both of which are shown to cope reasonably well with the physical principles of the protocol. The stability (ergodicity) conditions are obtained and interpreted in terms of maximum throughput. Several approximations related to these models are also discussed.

7.4.2.2. Model and optimization of CSMA

We have studied the maximum throughput of CSMA in scenarios with spatial reuse. The nodes of our network will be a Poisson Point Process (PPP) of a one or two dimensional space. The one dimensional well fits VANETs. To model the effect of Carrier Sense Multiple Access (CSMA), we give random marks to our nodes and to elect transmitting nodes in the PPP we choose the nodes with the smallest marks in their neighborhood, this is the Matern hardcore selection process. To describe the signal propagation, we use a signal with power-law decay and we add a random Rayleigh fading. To decide whether or not a transmission is successful, we adopt the Signal-over-Interference Ratio (SIR) model in which a packet is correctly received if its transmission power divided by the interference power is above a capture threshold. We assume that each node in our PPP has a random receiver at a typical distance. We choose the average distance to its closest neighbor. We also assume that all the network nodes always have a pending packet. With all these assumptions, we analytically study the density of throughput of successful transmission and we show that it can be optimized with the carrier-sense threshold.

7.4.2.3. Performance analysis of IEEE 802.11 broadcast schemes

We have analyzed different broadcast strategies in IEEE 802.11p Vehicular Ad-hoc NETWORKS (VANETs). The first strategy is the default IEEE 802.11p strategy. Using a model derived from the Bianchi model, we provide the network performance in terms of throughput and success rate. The second strategy is to use an

acknowledgment technique similar to the acknowledgment with point-to-point traffic. A node will send its broadcast packet as in the default case, but it requires an acknowledgment from a neighbor node. This node may be a random neighbor or may be selected according to precise rules. We analyze this second strategy in terms of throughput and success rate. Somewhat surprisingly, we show that this second strategy improves the delivery ratio of the transmitted packets but reduces the overall throughput. This means that if the CAM messages (Cooperative Awareness Messages) are broadcasted, the total number of packets actually delivered will be greater with the default strategy than with the improved strategy. We propose a third strategy which consists in using the default strategy for normal packets, but we add random redundant transmissions to ensure greater reliability for very important packets. We show that with this simple technique, not only do we obtain suitable reliability, but we also achieve larger global throughput than with the acknowledgment-oriented technique. We have also computed network performance in terms of throughput and success rate with respect to the network parameters and to analyze their impact on performances.

7.5. Models for wireless networks

7.5.1. *Interference and SINR coverage in spatial non-slotted Aloha networks*

Participants: Bartek Blaszczyszyn, Paul Muhlethaler.

We propose two analytically tractable stochastic-geometric models of interference in ad-hoc networks using pure (non-slotted) Aloha as the medium access. In contrast to the slotted model, the interference in pure Aloha may vary during the transmission of a tagged packet. We develop closed form expressions for the Laplace transform of the empirical average of the interference experienced during the transmission of a typical packet. Both models assume a power-law path-loss function with arbitrarily distributed fading and feature configurations of transmitters randomly located in the Euclidean plane according to a Poisson point process. Depending on the model, these configurations vary over time or are static. We apply our analysis of the interference to study the Signal-to-Interference-and-Noise Ratio (SINR) outage probability for a typical transmission in pure Aloha. The results are used to compare the performance of non-slotted Aloha to the slotted one, which has almost exclusively been previously studied in context of wired ad-hoc networks.

7.5.2. *Random linear multihop relaying in a general field of interferers using spatial Aloha*

Participants: Bartek Blaszczyszyn, Paul Muhlethaler.

We study a stationary Poisson pattern of nodes on a line embedded in an independent planar Poisson field of interfering nodes. Assuming slotted Aloha and the signal-to-interference-and-noise ratio capture condition, with the usual power-law path loss model and Rayleigh fading, we explicitly evaluate several local and end-to-end performance characteristics related to the nearest-neighbor packet relaying on this line, and study their dependence on the model parameters (the density of relaying and interfering nodes, Aloha tuning and the external noise power). Our model can be applied in two cases: the first use is for vehicular ad-hoc networks, where vehicles are randomly located on a straight road. The second use is to study a “typical” route traced in a (general) planar ad-hoc network by some routing mechanism. The approach we have chosen allows us to quantify the non-efficiency of long-distance routing in “pure ad-hoc” networks and evaluate a possible remedy for it in the form of additional “fixed” relaying nodes, called road-side units in a vehicular network. It also allows us to consider a more general field of interfering nodes and study the impact of the clustering of its nodes on the routing performance. As a special case of a field with more clustering than the Poisson field, we consider a Poisson-line field of interfering nodes, in which all the nodes are randomly located on random straight lines. In this case our analysis rigorously (in the sense of Palm theory) corresponds to the typical route of this network. The comparison to our basic model reveals a paradox: clustering of interfering nodes decreases the outage probability of a single (typical) transmission on the route, but increases the mean end-to-end delay

FOCUS Project-Team

7. New Results

7.1. Service-oriented computing

Participants: Maurizio Gabbrielli, Elena Giachino, Saverio Giallorenzo, Claudio Guidi, Mario Bravetti, Cosimo Laneve, Ivan Lanese, Michael Lienhardt, Jacopo Mauro, Fabrizio Montesi, Gianluigi Zavattaro.

7.1.1. Orchestrations

Orchestration models and languages in the context of Service-Oriented Architectures (SOA) are used to describe the composition of services focusing on their interactions. Concrete web services are connected to abstract service definitions for the aim of service discovery. In [16] we study a natural notion of compliance between clients and services in terms of their *bpel* (abstract) descriptions. The induced preorder shows interesting connections with the *must* preorder and has normal form representatives that are parallel-free finite-state activities, called *contracts*. Moreover, in [22] we focus on advancements of the orchestration language *Jolie* aiming at the development of dynamically adaptable orchestrated systems.

7.1.2. Choreographies

Choreographies are high-level descriptions of distributed interacting systems featuring as basic unit a communication between two participants. A main feature of choreographies is that they ensure deadlock-freedom by construction. From a choreography one can automatically derive a description of the behaviour of each participant using a notion of projection. Choreographies can be used both at the level of types (multiparty session types) or as a programming language. In [19] we surveyed our results about verification of adaptable processes, focusing in particular on distributed adaptability, where a process can update part of a protocol (specified by a choreography) by performing dynamic distributed updates over a set of protocol participants. In [14] we illustrate our approach to develop and verify distributed, adaptive software systems. The cornerstone of our framework is the use of choreography languages, which allow us to obtain correctness by construction. Moreover, in [36] we present *DIOC*, a language for programming distributed applications that are free from deadlocks and races by construction. A *DIOC* program describes a whole distributed application as a unique entity (choreography). *DIOC* allows the programmer to specify which parts of the application can be updated. At runtime, these parts may be replaced by new *DIOC* fragments from outside the application. *DIOC* programs are compiled, generating code for each site, in a lower-level language called *DPOC*. As a consequence *DPOC* applications are free from communication deadlocks and races, even in presence of runtime updates.

7.2. Models for reliability

Participants: Elena Giachino, Ivan Lanese.

7.2.1. Reversibility

We have continued the study of causal-consistent reversibility started in the past years. In [42] we defined the causal-consistent reversible semantics (both controlled and uncontrolled) of *muKlaim*, a formal coordination language based on distributed tuple spaces, by adapting the approach developed for message passing calculi in the past years. A major novelty is that the *muKlaim* read primitive allows two processes to access a shared resource independently, giving rise to a causality structure which is not found in message passing calculi.

In [31] we studied the issue of compliance of a client w.r.t. a server in a reversible setting using behavioural contracts. In particular, when an agreement cannot be reached, the client and the server can synchronously rollback to the last point of choice, looking for alternatives. As a main result, we showed that compliance is decidable even for recursive contracts.

7.3. Cloud Computing and Deployment

Participants: Elena Giachino, Saverio Giallorenzo, Claudio Guidi, Cosimo Laneve, Michael Lienhardt, Jacopo Mauro, Gianluigi Zavattaro.

7.3.1. Cloud application deployment

Configuration and management of applications in the cloud is a complex task that requires advanced methodologies and tools. A foundational study of the problem has been carried out in [21] where we have identified the critical tasks to be solved, quantified their computational complexity, and proposed simplifications to the problem with the idea of limiting the computational complexity at the cost of having approximated (but acceptable, in most cases) solutions. Our attention has been dedicated to the implementation of a tool for the efficient solution of one of these tasks, namely, the automatic planning of the management actions needed to properly configure a cloud application [17]. This tool, called Metis, has been already exploited in the realization of an integrated platform for the automatic deployment of the cloud application called Blender [39] as well as in the context of the ABS modeling language [37] in order to be able to support the automatic reasoning about deployment costs already during the early phases of application design and development. We have also performed a foundational study of the problem of reconfiguring an application instead of deploying it from scratch. Our foundational study allowed us, on the one hand, to quantify the computational complexity of the problem (PSpace-Complete) and, on the other hand, to precisely identify the source of such complexity (the presence of legacy components that cannot be re-deployed from scratch).

7.3.2. Cloud resource management

One of the main challenges in the management of cloud applications is the quantification of the computing resources needed by the applications to be deployed. More precisely, it is important to quantify upper bounds to the number of needed computing resources in order to either previously acquire them or have a precise quantification of the costs for executing an application. In [40] a static analysis technique is proposed that computes upper bounds of virtual machine usages in a concurrent language with explicit acquire and release operations of virtual machines. See the section on deadlock analysis for more details.

7.4. Resource Control and Probabilities

Participants: Michele Alberti, Martin Avanzini, Flavien Breuvert, Alberto Cappai, Ugo Dal Lago, Simone Martini, Giulio Pellitta, Alessandro Rioli, Davide Sangiorgi, Marco Solieri, Valeria Vignudelli.

7.4.1. Resource Control

7.4.1.1. Time Complexity Analysis of Concurrent and Higher-Order Functional Programs

We have extensively studied the problem of automatically analysing the complexity of programs. We first of all studied the problem for concurrent object-oriented programs [41]. To determine this complexity we have used intermediate abstract descriptions that record relevant information for the time analysis, called behavioural types. Behavioural types are then translated into so-called cost equations, making parallelism explicit. Cost equations are finally fed into an automatic off-the-shelf solver for obtaining the actual time complexity. The same problem has been also analysed when the underlying program is functional [29]. We showed how the complexity of higher-order functional programs can be analysed automatically by applying program transformations to a defunctionalized version of them, and feeding the result to existing tools for the complexity analysis of first-order term rewrite systems. This is done while carefully analysing complexity preservation and reflection of the employed transformations such that the complexity of the obtained term rewrite system reflects on the complexity of the initial program. This approach turns out to work well in practice, in particular since off-the-shelf complexity tool for first-order rewrite systems matured to a state where they are both fast and powerful. However, the implementation of such tools is quite sophisticated. To ensure correctness of the obtained complexity bounds, we extended CeTA, a certified proof checker for rewrite tools, with the formalisation of various complexity techniques underlying state-of-the-art complexity tools [30]. This way, we detected conflicts in theoretical results as well as bugs in existing complexity provers.

7.4.1.2. Function Algebras and Implicit Complexity

A fundamental result about ramified recurrence, one of the earliest systems in implicit complexity, has been proved [28]. This has been obtained through a careful analysis on how the adoption of an evaluation mechanism with sharing and memoization impacts the class of functions which can be computed in polynomial time. We have first shown how a natural cost model in which lookup for an already computed result has no cost is indeed invariant. As a corollary, we have then proved that the most general notion of ramified recurrence is sound for polynomial time.

7.4.1.3. Geometry of Interaction

We see the the geometry of interaction as a foundational framework in which the efficiency of higher-order computation can be analyzed. This has produced some very interesting results, also stimulated by the bilateral Inria project CRECOGI, which started this year. We have first of all studied the geometry of interaction of the resource lambda-calculus, a model of linear and nondeterministic functional languages. In a strictly typed restriction of the resource lambda-calculus, we have studied the notion of path persistence, and defined a geometry of interaction that characterises it [18]. Furthermore, we have carried out our work on multitoken machines, started in 2014. More specifically, we have studied multitoken interaction machines in the context of a very expressive linear logical system with exponentials, fixpoints and synchronization [34]. On the one hand, we have proved that interaction is guaranteed to be deadlock-free. On the other hand, the resulting logical system has been proved to be powerful enough to embed PCF and to adequately model its behaviour, both when call-by-name and when call-by-value evaluation are considered.

7.4.2. Probabilistic Models

7.4.2.1. Applicative Bisimilarity

Notions of equivalences for probabilistic programming languages have been studied and analysed, together with their relationships with context equivalence. More specifically, we have studied how applicative bisimilarity behaves when instantiated on a call-by-value probabilistic lambda-calculus, endowed with Plotkin's parallel disjunction operator [20]. We have proved that congruence and coincidence with the corresponding context relation hold for both bisimilarity and similarity, the latter known to be impossible in sequential languages. We have also shown that applicative bisimilarity works well when the underlying language of programs takes the form of a linear lambda-calculus extended with quantum data [35]. The main results are proofs of soundness for the obtained notion of equivalence.

7.4.2.2. From Equivalences to Metrics

The presence of probabilistic (thus quantitative) notions of observation makes equivalence relations too coarse-grained as ways to compare programs. This opens the way to metrics in which, indeed, not all non-equivalent programs are at the same distance. We have studied the problem of evaluating the distance between affine lambda-terms [33]. A natural generalisation of context equivalence has been shown to be characterised by a notion of trace distance, and to be bounded from above by a coinductively defined distance based on the Kantorovich metric on distributions. A different, again fully-abstract, tuple-based notion of trace distance has been shown to be able to handle nontrivial examples. A similar thing has been done in a calculus for probabilistic polynomial time computation [32], thus paving the way towards getting effective proof methodologies for computational indistinguishability, a key notion in modern cryptography.

7.5. Verification techniques for extensional properties

Participants: Daniel Hirschhoff, Elena Giachino, Michael Lienhardt, Cosimo Laneve, Jean-Marie Madiot, Davide Sangiorgi.

Extensional properties are those properties that constrain the behavioural descriptions of a system (i.e., how a system looks like from the outside). Examples of such properties include classical functional correctness, deadlock freedom and resource usage. Related to techniques for extensional properties are the issues of decidability (the problem of establishing whether certain properties are computationally feasible).

7.5.1. Static analysis of deadlock freedom and resource usage

Deadlock detection in concurrent programs that create networks with an arbitrary number of nodes is extremely complex and solutions either give imprecise answers or do not scale. To enable the analysis of such programs, we have studied an algorithm for detecting deadlocks in a basic concurrent object-oriented language. The algorithm (i) associates behavioural types, called *lam*, to programs by means of a type inference system and (ii) uses an ad-hoc verification technique highlighting circular dependencies in *lam* [15]. The algorithm has been prototyped and has been extended to a full-fledged programming language, called ABS.

A technique similar to [15] has been used for computing upper bounds of resource usages in [40]. In particular, the metaphor in this paper has been virtual machines usage in a concurrent language with explicit acquire and release operations. The problematic issue in such languages is when the release is delegated to other (ad-hoc or third party) concurrent codes (by passing them as arguments of invocations) – a feature that is currently used in Amazon Elastic Cloud Computing or in the Docker FiWare. As for deadlock analysis, the technique is modular and consists of (i) a type system associating programs with behavioural types that records relevant information for resource usage (creations, releases, and concurrent operations), (ii) a translation function that takes behavioural types and returns cost equations, and (iii) an automatic off-the-shelf solver for the cost equations. A soundness proof of the type system establishes the correctness of the above technique with respect to the cost equations. The technique has also been experimentally evaluated and the experiments show that it allows us to derive bounds for programs that are better than other techniques, such as those based on amortized analysis.

Another technique for enforcing program correctness is the one used in [36], [14], where the programming of distributed applications is guaranteed to be free from communication deadlocks and races by means of *choreographies*. Choreographies are behavioural types which allow one to obtain correctness by construction (more details on this paper in Section 7.1).

7.5.2. Name mobility

The article [44] studies the behavioural theory of $\pi\mathcal{P}$, a π -calculus featuring restriction as the only binder. In contrast with calculi such as Fusions and Chi, reduction in $\pi\mathcal{P}$ generates a preorder on names rather than an equivalence relation. Two characterisations of barbed congruence in $\pi\mathcal{P}$ are analyzed: the first is based on a compositional LTS, and the second is an axiomatisation. The results in this paper bring out basic properties of $\pi\mathcal{P}$, mostly related to the interplay between the restriction operator and the preorder on names.

7.5.3. Coinductive techniques

Coinductive techniques, notably those based on bisimulation, are widely used in concurrency theory to reason about systems of processes. The bisimulation proof method can be enhanced by employing “bisimulations up-to” techniques. A comprehensive theory of such enhancements has been developed for first-order (i.e., CCS-like) LTSs and bisimilarity, based on the notion of compatible function for fixed-point theory.

A proof method different from bisimulation is investigated in [46], [23]. This method is based on unique solution of special forms of inequations called contractions, and inspired by Milner’s theorem on unique solution of equations. The method is as powerful as the bisimulation proof method and its “up-to context” enhancements. The definition of contraction can be transferred onto other behavioural equivalences, possibly contextual and non-coinductive. This enables a coinductive reasoning style on such equivalences, either by applying the method based on unique solution of contractions, or by injecting appropriate contraction pre-orders into the bisimulation game. The technique can be applied both to first-order languages and to higher-order languages.

7.5.4. Expressiveness and decidability in actor-like systems

In [48], the limit of classical Petri nets is studied by discussing when it is necessary to move to the so-called Transfer nets, in which transitions can also move to a target place all the tokens currently present in a source place. More precisely, we consider a simple calculus of processes that interact by generating/consuming messages into/from a shared repository. For this calculus classical Petri nets can faithfully model the process

behavior. Then we present a simple extension with a primitive allowing processes to atomically rename all the data of a given kind. We show that with the addition of such primitive it is necessary to move to Transfer nets to obtain a faithful modeling.

7.6. Constraint Programming

Participants: Roberto Amadini, Maurizio Gabbrielli, Jacopo Mauro.

The Constraint Programming (CP) paradigm is a general and powerful framework that enables to express relations between different entities in form of constraints that must be satisfied. The concept of constraint is ubiquitous and not confined to the sciences: constraints appear in every aspect of daily life in the form of requirements, obligations, or prohibitions. Historically, the FOCUS group has always had an interest in CP, see e.g., [53], [54]. The possible applications of CP are in fact numerous and disparate. As an example, CP can be used for the deployment of services in the cloud [21], [39].

CP essentially consists of two layers: (i) a modeling level, in which a real-life problem is identified, examined, and formalized into a mathematical model by human experts; (ii) a solving level, aimed at resolving as efficiently and comprehensively as possible the model defined in (i) by means of software agents called constraint solvers. Over the last years we dealt with a particular aspect of CP, that is, the so called portfolio approaches [12], [27], [10]. In a nutshell, a portfolio approach in CP can be seen as the problem of predicting which is (are) the best constraint solver(s) —among a portfolio of available solvers— for solving a given CP problem. A constraint solver that relies on a portfolio of underlying, individual solvers is also dubbed a portfolio solver.

Our studies on portfolio approaches lead to development of the SUNNY-CP portfolio solver [26], [25]. SUNNY-CP relies on underlying state-of-the-art constraint solvers for solving a given CP problem encoded in the MiniZinc language, nowadays a de-facto standard for modeling CP problems. Initially developed as a sequential solver [26], SUNNY-CP has been later on enhanced by enabling the simultaneous execution of its solvers on different cores [25]. This extension allowed SUNNY-CP to win the gold medal in the open track of 2015 MiniZinc Challenge [cite], the annual competition for CP solvers.

However, we did not restrict the work on portfolio approaches to the CP field only. Indeed, we also performed some preliminary studies for evaluating SUNNY (i.e., the algorithm on which SUNNY-CP relies) in other application domains like, e.g., Boolean satisfiability (SAT), Quantified Boolean Formula (QBF), and Answer-Set Programming (ASP) [47], [24].

FUN Project-Team

7. New Results

7.1. Data gathering and coverage in WSN

Participants: Nathalie Mitton, Tahiry Razafindralambo, Arunabha Sen.

Data availability is one of the main goals and challenges in Future Ubiquitous Network and especially in Wireless Sensor Networks. Indeed, gathering and collecting data in a mobile environment is a very challenging task. In [12], the authors use data mules to organize the data collection from the sensor in the field. The results presented in [12] are based on some new and unique assumptions. First, it is assumed that the mules are mobile but also the sensors that generate the data to be collected. Second, we collection time is not the first optimization criteria. The focus of the paper is on minimizing the number of mules given a time constraints. The problem is shown to be NP complete and a transformation of the problem into a minimum flow problem allow the computation of optimal solution using Integer Linear Programming.

The results presented in [5] use the assumption of a mobile objects tracked by some other mobile objects as in [12]. In the case of [5] the focus is on coverage of mobile targets by mobile Unmanned Aerial Vehicle. The paper takes two major assumptions regarding the limited energy of UAV and the observation range. These two constraints are linked with each other since when the UAV increase its altitude, it consumes more energy but also increase its observation range. The problem under consideration is mathematically represented by defining mixed integer non-linear optimization models. Heuristic procedures are defined and they are based on restricted mixed integer programming (MIP) formulation of the problem. A computational study is carried out to assess the behavior of the proposed models and MIP-based heuristics.

7.2. Routing in FUN

Participants: Nathalie Mitton, Mouna Rekik.

Geographic routing is an attractive routing strategy in wireless sensor networks. It works well in dense networks, but it may suffer from the void problem. For this purpose, a recovery step is required to guarantee packet delivery. Face routing has widely been used as a recovery strategy since proved to guarantee delivery. However, it relies on a planar graph not always achievable in realistic wireless networks and may generate long paths. In [25], we propose GRACO, a new geographic routing algorithm that combines a greedy forwarding and a recovery strategy based on swarm intelligence. During recovery, ant packets search for alternative paths and drop pheromone trails to guide next packets within the network. GRACO avoids holes and produces near optimal paths. Simulation results demonstrate that GRACO leads to a significant improvement of routing performance and scalability when compared to the literature algorithms.

GRACO has first been designed in the general case. We then studied its applicability to the Virtual Power Plants and their specific data packets with different priorities [22], [24]. Indeed, the Smart Grid (SG) incorporates communication networks to the conventional electricity system in order to intelligently integrate distributed energy resources (DERs) and allow for demand side management. The move to Smart grid in developing countries has to cope with great disparities of ICT infrastructures even within the same city. Besides, individual DERs are often too small to be allowed access to energy market, likewise power utilities are unable to effectively control and manage small DERs. We propose the use of affordable and scalable wireless communication technology to aggregate geographically sparse DERs into a single virtual power plant. The enrollment of prosumers in the VPP is conditional to financial performance of the plant. Thus, the VPPs are dynamic and are expected to scale up as more and more prosumers are attracted by their financial benefits. the communication network has to follow this progression and therefore to be scalable and rapidly deploy-able. We present a routing algorithm for data communication within the VPP to support centralized, decentralized or fully distributed control of the VPP's DERs.

Based on this study, we adapted GRACO so it can fit the specific cases of Smart Grid [23] and more specifically to the Neighbor Area Networks (NAN) of Smart Grids, or distribution segment of the power system in the smart grid (SG). The deployment of ICT to support conventional grid will solve legacy problems that used to prevent implementation of smart services such as smart metering, demand side management or the integration of Distributed Energy Resources (DERs) within the smart grid. We demonstrate the effectiveness of GRACO in terms of scalability, peer-to-peer routing, end-to-end delay and delivery rate.

7.3. Deployment and Self-Deployment in FUN

Participants: Nathalie Mitton, Valeria Loscri, Tahiry Razafindralambo.

Mobility management is a difficult task in autonomous networks. However, mobility provide a huge advantage when in comes to specific scenario such as emergency-related ones especially when network connection must be restored to provide basic network access to users. [3] investigates the potential of spontaneous networks for providing Internet connectivity over the emergency area through the sharing of resources owned by the end-user devices. Novel and extremely flexible network deployment strategies are required in order to cope with the user mobility, the limited communication capabilities of wireless devices, and the intrinsic dynamics of traffic loads and QoS requirements. In [3], a novel architecture is proposed to take advantage of existing end-user devices and some algorithm, are described to build and efficiently exploit the spontaneous emergency network.

Following the emergency scenario described in [3], [15] and [21] describe an algorithm to minimizes the control traffic generated by specific nodes in the network used repair the network and the deployment of these specific nodes, This nodes, forming a substitution network, in case of emergency, are injected autonomously in the network by the network to restore basic network service. In order to increase the performance of the network, the injected nodes called substitution routers, use their ability to move to change the shape of the network and to increase its performance. These movements needs huge amount of control messages to maintain consistency regarding routers' positions. [15] and [21] give an algorithm for the deployment of these routers and the autoregressive time serie model to reduce the amount of control traffic used for the deployment.

7.4. Smart Cities

Participants: Nathalie Mitton, Valeria Loscri, Riccardo Petrolo.

Smart City represents one of the most promising, prominent and challenging Internet of Things (IoT) applications, but recent ICT trends suggest more and more that cities could also benefit from Cloud computing. The convergence of IoT paradigm and Cloud computing technology, can play a fundamental role for developing of highly level and organized cities form an ICT point of view, but it is of paramount importance to deal a critical analysis to identify the issues and challenges deriving from this synergy. This detailed study has been dealt in [7], where it is shown as the semantic annotation of the sensors in the cloud, and innovative services can be implemented and considered by bridging Cloud and Internet of Things. The Cloud of Things (CoT) paradigm is also considered in [16], where it is shown how the CoT arrives to better distribute resources, putting together and enabling therefore a horizontal integration of various Internet of Things(IoT) platforms. Semantic interoperability of diverse IoT platforms are also a key concept in [18], where the virtualization of different IoT systems in order to model and represent the architecture in accordance with the common standards-based IoT ontologies is applied. The environment comes with a range of visual drag-and-drop tools, which boosts developers' productivity.

7.5. RFID

Participant: Nathalie Mitton.

One of the devices under consideration by the FUN team is RFID. One of the main issues to widely deploy RFID reader is reader-to-reader collision. Indeed, when the electromagnetic fields of the readers overlap, a collision occurs on the tag laying in the overlapping section and cannot be read. In [10], we propose a high adaptive contention-based medium access control (HAMAC) protocol that considerably reduces readers collision problems in a large-scale dynamic RFID system. HAMAC is based only on realistic assumptions that can be experimented and does not require any additional components on RFID reader in order to improve the performance in terms of throughput, fairness and latency. The central idea of the HAMAC is for the RFID reader to use a WSN-like CSMA approach and to set its initial backoff counter to the maximum value that allows to mitigate collision. Then, according to the network congestion on physical channels the reader tries to dynamically control its contention window by linear decreasing on selected physical channel or multiplicative decreasing after scanning all available physical channels. Extensive simulations are proposed to highlight the performance of HAMAC compared to literature's work in large-scale RFID systems where both readers and tags are mobile. Simulation results show the effectiveness and robustness of the proposed anti-collision protocol in terms of network throughput, fairness, coverage and time to read all tags.

7.6. Localization

Participants: Nathalie Mitton, Roudy Dagher, Valeria Loscri, Salvatore Guzzo Bonifacio.

[20] presents our approach to localize a node with the use of only one landmark. It is a passive and non intrusive cross-layer approach that relies on a signal processing of all received signals. Results are evaluated by simulation and show good accuracy. To complete the previous study, we developed [11] a novel array-based method to estimate the path loss exponent (PLE). The method is designed as a part of an automatic calibration step, prior to localization of a source transmitting in the near-far field of the array. The method only requires the knowledge of the ranges between the array elements. By making the antenna elements transmit in turn, the array response model in the near-far field is exploited to estimate the current environment PLE. Simulation results show that this method can achieve good performance with one transmission round. The performance of the PLE estimation is investigated in the context of source localization with a sensitivity analysis to the PLE estimation. These works are the purpose of a pending patent (submitted in March 2015).

Alternatively, we derive similar localization schemes to enable a cooperation between mobile robots to locate a target based on RSSI [13]. Received Signal Strength Indicator (RSSI) is commonly considered and is very popular for target localization applications, since it does not require extra-circuitry and is always available on current devices. Unfortunately, target localizations based on RSSI are affected with many issues, above all in indoor environments. In this paper, we focus on the pervasive localization of target objects in an unknown environment. In order to accomplish the localization task, we implement an Associative Search Network (ASN) on the robots and we deploy a real test-bed to evaluate the effectiveness of the ASN for target localization. The ASN is based on the computation of weights, to "dictate" the correct direction of movement, closer to the target. Results show that RSSI through an ASN is effective to localize a target, since there is an implicit mechanism of correction, deriving from the learning approach implemented in the ASN.

7.7. Vehicular Networks

Participants: Nathalie Mitton, Valeria Loscri.

In the framework of our collaboration with Southern University in China, we investigate a specific issue in Vehicular AdHoc Networks (VANET), the information delivery delay analysis for roadside unit deployment in a VANET with intermittent connectivity [9]. A mathematical model is developed to describe the relationship between the average delay for delivering road condition information and the distance between two neighbor RSUs deployed along a road. The derived mathematical model considers a straight highway scenario where two RSUs are deployed at a distance without any direct connection and vehicles are sparsely distributed on the road with road condition information randomly generated between the two neighbor RSUs. Moreover, the model takes into account the vehicle speed, the vehicle density, the likelihood of an incident, and the distance between two RSUs. The effectiveness of the derived mathematical model is verified through simulation results.

Given the information delivery delay constraint of a time-critical application, this model can be used to estimate the maximum distance allowed between two neighbor RSUs, which can provide a reference for the deployment of RSUs in such scenarios.

But Vehicular Networks can also convey social networks. In [30], we survey recent literature on Vehicular Social Networks that are a particular class of vehicular ad hoc networks, characterized by social aspects and features. Starting from this pillar, we investigate perspectives of next generation vehicles under the assumption of social networking for vehicular applications (i.e., safety and entertainment applications). This paper plays a role as a starting point about socially-inspired vehicles, and main related applications, as well as communication techniques. Vehicular communications can be considered as the "first social network for automobiles", since each driver can share data with other neighbors. As an instance, heavy traffic is a common occurrence in some areas on the roads (e.g., at intersections, taxi loading/unloading areas, and so on); as a consequence, roads become a popular social place for vehicles to connect to each other. Human factors are then involved in vehicular ad hoc networks, not only due to the safety related applications, but also for entertainment purpose. Social characteristics and human behavior largely impact on vehicular ad hoc networks, and this arises to the vehicular social networks, which are formed when vehicles (individuals) "socialize" and share common interests. This survey describes the main features of vehicular social networks, from novel emerging technologies to social aspects used for mobile applications, as well as main issues and challenges. Vehicular social networks are described as decentralized opportunistic communication networks formed among vehicles. They exploit mobility aspects, and basics of traditional social networks, in order to create novel approaches of message exchange through the detection of dynamic social structures. An overview of the main state-of-the-art on safety and entertainment applications relying on social networking solutions is also provided.

7.8. FIT

Participants: Nathalie Mitton, Julien Vandaele.

The universal proliferation of intelligent objects is making Internet of Things (IoT) a reality; to operate on a large scale it will critically rely on new, seamless, forms of communications. But how can innovations be validated in a controlled environment, before being massively deployed into the real world? Several platforms have been deployed to address this issue. In [8], we browse a survey of them, highlighting their characteristics and given some tips to choose the most appropriate to our needs.

Our team has contributed to the deployment of the FIT IoT-LAB platform [2], [19], [27], which addresses this challenge by offering a unique open first class service to all IoT developers, researchers, integrators and developers: a large-scale experimental testbed allowing design, development, deployment and testing of innovative IoT applications, in order to test the future and make it safe. One of the specific deployment focuses on the automatic docking of robots for energy recharge. We explain it in [17]. The objective is to achieve long-term autonomous robots within an experiment test-bed. We propose to combine the use of QR codes as landmarks and Infrared distance sensors. The relative size of the lateral edges of the visual pattern is used to position the robot in relation with the dock. Infrared distance sensors are then used to perform different approaching strategies depending on the distance. Experiments show that the proposed solution is fully operational and robust. Not to rely exclusively on visual pattern recognition avoids potential errors induced by camera calibration. Additionally, as a positive side effect, the use of Infrared sensors allows the robot to avoid obstacles while docking. The finality of such an approach is to integrate these robots into the FIT IoT Lab experimental testbed which allows any experimenter to book wireless resources such as wireless sensors remotely and to test their own code. Wifibots holding wireless sensors will be integrated as additional reservable resources of the platform to enlarge the set of possible experimentations with mobile entities.

7.9. New and other communication paradigms

Participants: Nathalie Mitton, Valeria Loscri, Arash Maskooki, Gabriele Sabatino.

Interconnection and self-organized systems are normally populated with heterogeneous and different devices. The differences range from computational capabilities, storage size, etc. Instead of considering the heterogeneity as a limitation, it is possible to "turn it" as a primitive control of the system, in order to realize more robust and more resilient communication systems. Based on these premises, we identify specific situations, where mobile nodes with a plethora of interesting features and sensing capacities, can be exploited by configuring them in such a way to make them playing different roles in respect of them for which they have been initially conceived [4]. The differentiated use of devices, together with a careful analysis of the characteristics and performance requirements of the current and the future networks, allow the adaptation to the exponential growth in demand for high bandwidth applications [26]. This is exactly the philosophy embraced in [28], where Software Defined Radio (SDR) and Cognitive Radio (CR) have been considered and analyzed in a novel context, namely body networked systems. A detailed analysis of body systems as networked systems has also been considered in [6] and [14]. In [6] a novel communication paradigm, namely a molecular communication, has been considered to show how a nanoparticulate system can be suitable to coexist in a biological environment. An experimental analysis to assess the theoretical assumption has been developed in [14]. In order to assess new/alternative communication paradigms, there is the necessity from one side to consider and analyze the specific context and its level of interaction with the communication system and on the other side the correct identification of the specific features of the communication paradigm itself. This type of analysis allowed the design and implementation of an acoustic communication approach [29], where the ultrasound represent the wave carriers of data information. This "unusual" transmission means has been selected as the most suitable in a context as the body, where the aqueous environment makes it not suitable for more "traditional" communication paradigms, e.g. the one based on Radio Frequency (RF) waves.

GANG Project-Team

7. New Results

7.1. Graph and Combinatorial Algorithms

7.1.1. Rainbow matchings in hypergraphs

A rainbow matching for (not necessarily distinct) sets F_1, \dots, F_k of hypergraph edges is a matching consisting of k edges, one from each F_i . In [8], we give some order to the multitude of conjectures that relate to this concept, as well as introduce some new conjectures. We also present some partial results on one of these conjectures, that seems central among them – the so-called Ryser-Brauer-Stein conjecture.

7.1.2. A graph formulation of the union-closed sets conjecture

In 1979, Frankl conjectured that in a finite non-trivial union-closed collection of sets there has to be an element that belongs to at least half the sets. In [7], we show that this is equivalent to the conjecture that in a finite non-trivial graph there are two adjacent vertices, each belonging to at most half of the maximal stable sets. In this graph formulation other special cases become natural. The conjecture is trivially true for non-bipartite graphs and we show that it also holds for the classes of chordal bipartite graphs, subcubic bipartite graphs, bipartite series-parallel graphs and bipartitioned circular interval graphs.

7.1.3. Cops-and-robber games on k -chordal graphs

The cops-and-robber games, introduced by Winkler and Nowakowski (in Discrete Math. 43, 1983) and independently defined by Quilliot (in J. Comb. Theory, Ser. B 38, 1985), concern a team of cops that must capture a robber moving in a graph. In [20], we consider the class of k -chordal graphs, i.e., graphs with no induced (chordless) cycle of length greater than k , $k \geq 3$. We prove that $k-1$ cops are always sufficient to capture a robber in k -chordal graphs. This leads us to our main result, a new structural decomposition for a graph class including k -chordal graphs.

We present a polynomial-time algorithm that, given a graph G and $k \geq 3$, either returns an induced cycle larger than k in G , or computes a tree-decomposition of G , each bag of which contains a dominating path with at most $k-1$ vertices. This allows us to prove that any k -chordal graph with maximum degree Δ has treewidth at most $(k-1)(\Delta-1) + 2$, improving the $O(\Delta(\Delta-1)k-3)$ bound of Bodlaender and Thilikos (Discrete Appl. Math. 79, 1997). Moreover, any graph admitting such a tree-decomposition has small hyperbolicity). As an application, for any n -vertex graph admitting such a tree-decomposition, we propose a compact routing scheme using routing tables, addresses and headers of size $O(k \log \Delta + \log n)$ bits and achieving an additive stretch of $O(k \log \Delta)$. As far as we know, this is the first routing scheme with $O(k \log \Delta + \log n)$ -routing tables and small additive stretch for k -chordal graphs.

7.1.4. Distinguishing views in symmetric networks

The view of a node in a port-labeled network is an infinite tree encoding all walks in the network originating from this node. In [16], we prove that for any integers $n \geq D \geq 1$, there exists a port-labeled network with at most n nodes and diameter at most D , which contains a pair of nodes whose (infinite) views are different, but whose views truncated to depth $\Omega(D \log(n/D))$ are identical.

7.1.5. Vertex elimination orderings for hereditary graph classes

In [3], we provide a general method to prove the existence and compute efficiently elimination orderings in graphs. This method relies on several tools that were known before, but that were not put together so far: the algorithm LexBFS due to Rose, Tarjan and Lueker, its additional properties discovered by Berry and Bordat, and a local decomposition property of graphs discovered by Maffray, Trotignon and Vušković. We use this method to prove the existence of elimination orderings in several classes of graphs, and to compute them in linear time. Some of the classes have already been studied, namely even-hole-free graphs, square-theta-free Berge graphs, universally signable graphs and wheel-free graphs. Some other classes are new. It turns out that all the classes that we consider can be defined by excluding some of the so-called Truemper configurations. For several classes of graphs, we obtain directly bounds on the chromatic number, or fast algorithms for the maximum clique problem or the coloring problem.

7.1.6. Fast collaborative graph exploration

In [14], we study the following scenario of online graph exploration. A team of k agents is initially located at a distinguished vertex r of an undirected graph. We ask how many time steps are required to complete exploration, i.e., to make sure that every vertex has been visited by some agent. As our main result, we provide the first strategy which performs exploration of a graph with n vertices at a distance of at most D from r in time $O(D)$, using a team of agents of polynomial size $k = Dn^{1+\epsilon} < n^{2+\epsilon}$, for any $\epsilon > 0$. Our strategy works in the local communication model, in which agents can only exchange information when located at a vertex, without knowledge of global parameters such as n or D .

We also obtain almost-tight bounds on the asymptotic relation between exploration time and team size, for large k , in both the local and the global communication model.

7.1.7. Position discovery for a system of bouncing robots

In [11], we consider a scenario in which a collection of n anonymous mobile robots is deployed on a unit-perimeter ring or a unit-length line segment. Every robot starts moving at constant speed, and bounces each time it meets any other robot or segment endpoint, changing its walk direction. We study the problem of position discovery, in which the task of each robot is to detect the presence and the initial positions of all other robots. The robots cannot communicate or perceive information about the environment in any way other than by bouncing nor they have control over their walks which are determined by their initial positions and their starting directions. Each robot has a clock allowing it to observe the times of its bounces. We give complete characterizations of all initial configurations for both the ring and the segment in which no position detection algorithm exists and we design optimal position detection algorithms for all feasible configurations.

7.1.8. Rendezvous of mobile agents in edge-weighted networks

In [15], we introduce a variant of the deterministic rendezvous problem for a pair of heterogeneous agents operating in an undirected graph, which differ in the time they require to traverse particular edges of the graph. Each agent knows the complete topology of the graph and the initial positions of both agents. The agent also knows its own traversal times for all of the edges of the graph, but is unaware of the corresponding traversal times for the other agent. The goal of the agents is to meet on an edge or a node of the graph. In this scenario, we study the time required by the agents to meet, compared to the meeting time T_{OPT} in the offline scenario in which the agents have complete knowledge about each others' speed characteristics. When no additional assumptions are made, we show that rendezvous in our model can be achieved after time $O(nT_{OPT})$ in a n -node graph, and that such time is essentially in some cases the best possible. However, we prove that the rendezvous time can be reduced to $\Theta(T_{OPT})$ when the agents are allowed to exchange $\Theta(n)$ bits of information at the start of the rendezvous process. We then show that under some natural assumption about the traversal times of edges, the hardness of the heterogeneous rendezvous problem can be substantially decreased, both in terms of time required for rendezvous without communication, and the communication complexity of achieving rendezvous in time $\Theta(T_{OPT})$.

7.1.9. Monitoring a graph using faulty mobile robots

In the scenario studied in [27], a team of k mobile robots is deployed on a weighted graph whose edge weights represent distances. The robots perpetually move along the domain, represented by all points belonging to the graph edges, not exceeding their maximal speed. The robots need to patrol the graph by regularly visiting all points of the domain. Here, we consider a team of robots (patrolmen), at most f of which may be unreliable, i.e. they fail to comply with their patrolling duties.

What algorithm should be followed so as to minimize the maximum time between successive visits of every edge point by a reliable patrolmen? The corresponding measure of efficiency of patrolling called idleness has been widely accepted in the robotics literature. We extend it to the case of untrusted patrolmen; we denote by $I_k^f(G)$ the maximum time that a point of the domain may remain unvisited by reliable patrolmen. The objective is to find patrolling strategies minimizing $I_k^f(G)$.

We investigate this problem for various classes of graphs. We design optimal algorithms for line segments, which turn out to be surprisingly different from strategies for related patrolling problems proposed in the literature. We then use these results to study the case of general graphs. For Eulerian graphs G , we give an optimal patrolling strategy with idleness $I_k^f(G) = (f + 1)|E|/k$, where $|E|$ is the sum of the lengths of the edges of G . Further, we show the hardness of the problem of computing the idle time for three robots, at most one of which is faulty, by reduction from 3-edge-coloring of cubic graphs — a known NP-hard problem. A byproduct of our proof is the investigation of classes of graphs minimizing idle time (with respect to the total length of edges); an example of such a class is known in the literature under the name of Kotzig graphs.

7.1.10. Limit behavior of the rotor-router system

The rotor-router model, also called the Propp machine, was introduced as a deterministic alternative to the random walk. In this model, a group of identical tokens are initially placed at nodes of the graph. Each node maintains a cyclic ordering of the outgoing arcs, and during consecutive turns the tokens are propagated along arcs chosen according to this ordering in round-robin fashion. The behavior of the model is fully deterministic. Yanovski et al. (Algorithmica, 2003) proved that a single rotor-router walk on any graph with m edges and diameter D stabilizes to a traversal of an Eulerian circuit on the set of all $2m$ directed arcs on the edge set of the graph, and that such periodic behaviour of the system is achieved after an initial transient phase of at most $2mD$ steps.

The case of multiple parallel rotor-routers was studied experimentally, leading Yanovski et al. to the experimental observation that a system of $k > 1$ parallel walks also stabilizes with a period of length at most $2m$ steps. In our work [26] we disprove this observation, showing that the period of parallel rotor-router walks can in fact, be superpolynomial in the size of graph. On the positive side, we provide a characterization of the periodic behavior of parallel router walks, in terms of a structural property of stable states called a subcycle decomposition. This property provides us the tools to efficiently detect whether a given system configuration corresponds to the transient or to the limit behavior of the system. Moreover, we provide polynomial upper bounds of $O(m^4D^2 + mD \log k)$ and $O(m^5k^2)$ on the number of steps it takes for the system to stabilize. Thus, we are able to predict any future behavior of the system using an algorithm that takes polynomial time and space. In addition, we show that there exists a separation between the stabilization time of the single-walk and multiple-walk rotor-router systems, and that for some graphs the latter can be asymptotically larger even for the case of $k = 2$ walks.

7.2. Distributed Computing

7.2.1. Self-stabilizing verification and computation of an MST

In the work [19], we demonstrate the usefulness of distributed local verification of proofs, as a tool for the design of self-stabilizing algorithms. In particular, it introduces a somewhat generalized notion of distributed local proofs, and utilizes it for improving the time complexity significantly, while maintaining space optimality. As a result, we show that optimizing the memory size carries at most a small cost in terms of time, in the context of Minimum Spanning Tree (MST). That is, we present algorithms that are both time and space efficient for

both constructing an MST and for verifying it. This involves several parts that may be considered contributions in themselves.

First, we generalize the notion of local proofs, trading off the time complexity for memory efficiency. This adds a dimension to the study of distributed local proofs, which has been gaining attention recently. Specifically, we design a (self-stabilizing) proof labeling scheme which is memory optimal (i.e., $O(\log n)$ bits per node), and whose time complexity is $O(\log^2 n)$ in synchronous networks, or $O(\Delta \log^3 n)$ time in asynchronous ones, where Δ is the maximum degree of nodes. This answers an open problem posed by Awerbuch and Varghese (FOCS 1991). We also show that $\Omega(\log n)$ time is necessary, even in synchronous networks. Another property is that if f faults occurred, then, within the required detection time above, they are detected by some node in the $O(f \log n)$ locality of each of the faults. Second, we show how to enhance a known transformer that makes input/output algorithms self-stabilizing. It now takes as input an efficient construction algorithm and an efficient self-stabilizing proof labeling scheme, and produces an efficient self-stabilizing algorithm. When used for MST, the transformer produces a memory optimal self-stabilizing algorithm, whose time complexity, namely, $O(n)$, is significantly better even than that of previous algorithms. (The time complexity of previous MST algorithms that used $O(\log^2 n)$ memory bits per node was $O(n^2)$, and the time for optimal space algorithms was $O(n|E|)$.) Inherited from our proof labelling scheme, our self-stabilising MST construction algorithm also has the following two properties: (1) if faults occur after the construction ended, then they are detected by some nodes within $O(\log^2 n)$ time in synchronous networks, or within $O(\Delta \log^3 n)$ time in asynchronous ones, and (2) if f faults occurred, then, within the required detection time above, they are detected within the $O(f \log n)$ locality of each of the faults. We also show how to improve the above two properties, at the expense of some increase in the memory.

7.2.2. Clock synchronization and distributed estimation in highly dynamic networks

In [21], we consider the External Clock Synchronization problem in dynamic sensor networks. Initially, sensors obtain inaccurate estimations of an external time reference and subsequently collaborate in order to synchronize their internal clocks with the external time. For simplicity, we adopt the drift-free assumption, where internal clocks are assumed to tick at the same pace. Hence, the problem is reduced to an estimation problem, in which the sensors need to estimate the initial external time. In this context of distributed estimation, this work is further relevant to the problem of collective approximation of environmental values by biological groups.

Unlike most works on clock synchronization that assume static networks, the setting considered here is an extreme case of highly dynamic networks. We do however impose a restriction on the dynamicity of the network. Specifically, we assume a non-adaptive scheduler adversary that dictates an arbitrary, yet independent, meeting pattern. Such meeting patterns fit, for example, with short-time scenarios in highly dynamic settings, where each sensor interacts with only few other arbitrary sensors.

We propose an extremely simple clock synchronization (or an estimation) algorithm that is based on weighted averages, and prove that its performance on any given independent meeting pattern is highly competitive with that of the best possible algorithm, which operates without any resource or computational restrictions, and further knows the whole meeting pattern in advance. In particular, when all distributions involved are Gaussian, the performances of our scheme coincide with the optimal performances. Our proofs rely on an extensive use of the concept of Fisher information. We use the Cramér-Rao bound and our definition of a Fisher Channel Capacity to quantify information flows and to obtain lower bounds on collective performance. This opens the door for further rigorous quantifications of information flows within collaborative sensors.

7.2.3. Wait-freedom with advice

In [13], we motivate and propose a new way of thinking about failure detectors which allows us to define what it means to solve a distributed task wait-free using a failure detector. In our model, the system is composed of computation processes that obtain inputs and are supposed to produce outputs and synchronization processes that are subject to failures and can query a failure detector. Under the condition that correct (never failing) synchronization processes take sufficiently many steps, they provide the computation processes with enough advice to solve the given task wait-free: every computation process outputs in a finite number of its own

steps, regardless of the behavior of other computation processes. Every task can thus be characterized by the weakest failure detector that allows for solving it, and we show that every such failure detector captures a form of set agreement. We then obtain a complete classification of tasks, including ones that evaded comprehensible characterization so far, such as renaming or weak symmetry breaking.

7.2.4. Linear-space bootstrap communication schemes

In [12], we consider a system of n processes with ids that are drawn from a large space. How can these n processes communicate to solve a problem? It is shown that linear number of Multi-Writer Multi-Reader (MWMR) registers are sufficient to solve any read-write wait-free solvable problem and needed to solve some read-write wait-free solvable problem. This contrasts with the existing possible solution borrowed from adaptive algorithms that require $\Theta(n^{3/2})$ MWMR registers.

To obtain the sufficiency result, we show how the processes can non-blockingly emulate a system of n Single-Writer Multi-Reader (SWMR) registers on top of n Multi-Writer Multi-Reader (MWMR) registers. For the necessity result, we show it is impossible to do such an emulation with $n-1$ MWMR registers.

We also presents a wait-free emulation, using $2n-1$ rather than just n registers. The emulation can be used to solve an infinite sequence of tasks that are sequentially dependent (processes need the previous task's outputs in order to proceed to the next task). A non-blocking emulation cannot be used in this case, because it might starve a process forever.

7.2.5. Space complexity of set agreement

The k -set agreement problem is a generalization of the classical consensus problem in which processes are permitted to output up to k different input values. In a system of n processes, an m -obstruction-free solution to the problem requires termination only in executions where the number of processes taking steps is eventually bounded by m . This family of progress conditions generalizes wait-freedom ($m = n$) and obstruction-freedom ($m = 1$). In [29], we prove upper and lower bounds on the number of registers required to solve m -obstruction-free k -set agreement, considering both one-shot and repeated formulations. In particular, we show that repeated k set agreement can be solved using $n + 2m - k$ registers and establish a nearly matching lower bound of $n + m - k$.

7.2.6. Consensus capability of distributed systems

A fundamental research theme in distributed computing is the comparison of systems in terms of their ability to solve basic problems such as consensus that cannot be solved in completely asynchronous systems. In particular, in a seminal work (ACM Trans. Program. Lang. Syst. 13, 1991), Herlihy compares shared-memory systems in terms of the shared objects that they have: he proved that there are shared objects that are powerful enough to solve consensus for n processes, but are too weak to solve consensus for $n + 1$ processes; such objects are placed at level n of a wait-free hierarchy.

Similarly as in that work, in [30] we compare shared-memory systems with respect to their ability to solve consensus for n processes. But instead of comparing systems defined by the shared objects that they have, we compare read-write systems defined by the set of process schedules that can occur in these systems. Defining systems this way can capture many types of systems, e.g., systems whose synchrony ranges from fully synchronous to completely asynchronous, several systems with failure detectors, and "obstruction-free" systems. Here, we consider read-write systems defined in terms of sets of process schedules, and investigate the following fundamental question: Is there a system of $n + 1$ processes such that consensus can be solved for every subset of n processes in the system, but consensus cannot be solved for the $n + 1$ processes of the system? We show that the answer to the above question is "yes", and so these systems can be classified into a hierarchy akin to Herlihy's hierarchy.

7.2.7. Shared whiteboard models of distributed systems

In [4], we study distributed algorithms on massive graphs where links represent a particular relationship between nodes (for instance, nodes may represent phone numbers and links may indicate telephone calls).

Since such graphs are massive they need to be processed in a distributed way. When computing graph-theoretic properties, nodes become natural units for distributed computation. Links do not necessarily represent communication channels between the computing units and therefore do not restrict the communication flow. Our goal is to model and analyze the computational power of such distributed systems where one computing unit is assigned to each node. Communication takes place on a whiteboard where each node is allowed to write at most one message. Every node can read the contents of the whiteboard and, when activated, can write one small message based on its local knowledge. When the protocol terminates its output is computed from the final contents of the whiteboard. We describe four synchronization models for accessing the whiteboard. We show that message size and synchronization power constitute two orthogonal hierarchies for these systems. We exhibit problems that separate these models, i.e., that can be solved in one model but not in a weaker one, even with increased message size. These problems are related to maximal independent set and connectivity. We also exhibit problems that require a given message size independently of the synchronization model.

7.2.8. Discrete Lotka-Volterra population protocols

In [28], we focus on a natural class of population protocols whose dynamics are modeled by the discrete version of Lotka-Volterra equations with no linear term. In such protocols, when an agent a of type (species) i interacts with an agent b of type (species) j with a as the initiator, then b 's type becomes i with probability P_{ij} . In such an interaction, we think of a as the predator, b as the prey, and the type of the prey is either converted to that of the predator or stays as is. Such protocols capture the dynamics of some opinion spreading models and generalize the well-known Rock-Paper-Scissors discrete dynamics. We consider the pairwise interactions among agents that are scheduled uniformly at random.

We start by considering the convergence time and show that any Lotka-Volterra-type protocol on a n -agent population converges to some absorbing state in time polynomial in n , w.h.p., when any pair of agents is allowed to interact. By contrast, when the interaction graph is a star, there exist protocols of the considered type, such as Rock-Paper-Scissors, which require exponential time to converge. We then study threshold effects exhibited by Lotka-Volterra-type protocols with 3 and more species under interactions between any pair of agents. We present a simple 4-type protocol in which the probability difference of reaching the two possible absorbing states is strongly amplified by the ratio of the initial populations of the two other types, which are transient, but “control” convergence. We then prove that the Rock-Paper-Scissors protocol reaches each of its three possible absorbing states with almost equal probability, starting from any configuration satisfying some sub-linear lower bound on the initial size of each species. That is, Rock-Paper-Scissors is a realization of a “coin-flip consensus” in a distributed system. Some of our techniques may be of independent value.

7.2.9. Deterministic load-balancing

In [23], we consider the problem of deterministic load balancing of tokens in the discrete model. A set of n processors is connected into a d -regular undirected network. In every time step, each processor exchanges some of its tokens with each of its neighbors in the network. The goal is to minimize the discrepancy between the number of tokens on the most-loaded and the least-loaded processor as quickly as possible. Rabani et al. (FOCS 1998) present a general technique for the analysis of a wide class of discrete load balancing algorithms. Their approach is to characterize the deviation between the actual loads of a discrete balancing algorithm with the distribution generated by a related Markov chain. The Markov chain can also be regarded as the underlying model of a continuous diffusion algorithm. Rabani et al. showed that after time $T = O(\log(Kn)/\mu)$, any algorithm of their class achieves a discrepancy of $O(d \log n/\mu)$, where μ is the spectral gap of the transition matrix of the graph, and K is the initial load discrepancy in the system.

In this work we identify some natural additional conditions on deterministic balancing algorithms, resulting in a class of algorithms reaching a smaller discrepancy. This class contains well-known algorithms, e.g., the rotor-router. Specifically, we introduce the notion of cumulatively fair load-balancing algorithms where in any interval of consecutive time steps, the total number of tokens sent out over an edge by a node is the same (up to constants) for all adjacent edges. We prove that algorithms which are cumulatively fair and where every node retains a sufficient part of its load in each step, achieve a discrepancy of $O(d\sqrt{\log n/\mu}, d\sqrt{n})$ in time $O(T)$. We also show that in general neither of these assumptions may be omitted without increasing discrepancy. We

then show by a combinatorial potential reduction argument that any cumulatively fair scheme satisfying some additional assumptions achieves a discrepancy of $O(d)$ almost as quickly as the continuous diffusion process. This positive result applies to some of the simplest and most natural discrete load balancing schemes.

7.2.10. Randomized local network computing

In [32], we have carried on investigating the line of research questioning the power of randomization for the design of distributed algorithms. In their seminal paper, Naor and Stockmeyer [STOC 1993] established that, in the context of network computing, in which all nodes execute the same algorithm in parallel, any construction task that can be solved locally by a randomized Monte-Carlo algorithm can also be solved locally by a deterministic algorithm. This result however holds in a specific context. In particular, it holds only for distributed tasks whose solutions that can be locally checked by a deterministic algorithm. We have extended the result of Naor and Stockmeyer to a wider class of tasks. Specifically, we proved that the same derandomization result holds for every task whose solutions can be locally checked using a 2-sided error randomized Monte-Carlo algorithm. This extension finds applications to, e.g., the design of lower bounds for construction tasks which tolerate that some nodes compute incorrect values. In a nutshell, we have showed that randomization does not help for solving such resilient tasks.

7.2.11. Proof-labeling schemes: randomization and self-stabilization

We have also carried on investigating the power of randomization for the design of proof-labeling schemes. Recall that a proof-labeling scheme, introduced by Korman, Kutten and Peleg [PODC 2005], is a mechanism enabling to certify the legality of a network configuration with respect to a boolean predicate. Such a mechanism finds applications in many frameworks, including the design of fault-tolerant distributed algorithms. In a proof-labeling scheme, the verification phase consists of exchanging labels between neighbors. The size of these labels depends on the network predicate to be checked. There are predicates requiring large labels, of poly-logarithmic size (e.g., MST), or even polynomial size (e.g., Symmetry). In [22], we introduce the notion of randomized proof-labeling schemes. By reduction from deterministic schemes, we show that randomization enables the amount of communication to be exponentially reduced. As a consequence, we show that checking any network predicate can be done with probability of correctness as close to one as desired by exchanging just a logarithmic number of bits between neighbors. Moreover, we design a novel space lower bound technique that applies to both deterministic and randomized proof-labeling schemes. Using this technique, we establish several tight bounds on the verification complexity of classical distributed computing problems, such as MST construction, and of classical predicates such as acyclicity, connectivity, and cycle length.

Next, we have established the formal connections between self-stabilization and proof-labeling scheme. Recall that self-stabilizing algorithms are distributed algorithms supporting transient failures. Starting from any configuration, they allow the system to detect whether the actual configuration is legal, and, if not, they allow the system to eventually reach a legal configuration. In the context of network computing, it is known that, for every task, there is a self-stabilizing algorithm solving that task, with optimal space-complexity, but converging in an exponential number of rounds. On the other hand, it is also known that, for every task, there is a self-stabilizing algorithm solving that task in a linear number of rounds, but with large space-complexity. It is however not known whether for every task there exists a self-stabilizing algorithm that is simultaneously space-efficient and time-efficient. In [24], we make a first attempt for answering the question of whether such an efficient algorithm exists for every task, by focussing on constrained spanning tree construction tasks. We present a general roadmap for the design of silent space-optimal self-stabilizing algorithms solving such tasks, converging in polynomially many rounds under the unfair scheduler. By applying our roadmap to the task of constructing minimum-weight spanning tree (MST), and to the task of constructing minimum-degree spanning tree (MDST), we provide algorithms that outperform previously known algorithms designed and optimized specifically for solving each of these two tasks.

7.2.12. Role of node identifiers in local decision

We have also investigated the role of IDs in network computing. This role is well understood as far as symmetry breaking is concerned. However, the unique identifiers also leak information about the computing environment

— in particular, they provide some nodes with information related to the size of the network. It was recently proved that in the context of local decision, there are some decision problems such that (1) they cannot be solved without unique identifiers, and (2) unique node identifiers leak a sufficient amount of information such that the problem becomes solvable. In [33] we study what is the minimal amount of information that we need to leak from the environment to the nodes in order to solve local decision problems. Our key results are related to scalar oracles f that, for any given n , provide a multi-set $f(n)$ of n labels; then the adversary assigns the labels to the n nodes in the network. This is a direct generalization of the usual assumption of unique node identifiers. We give a complete characterization of the weakest oracle that leaks at least as much information as the unique identifiers. Our main result is the following dichotomy: we classify scalar oracles as large and small, depending on their asymptotic behavior, and show that (1) any large oracle is at least as powerful as the unique identifiers in the context of local decision problems, while (2) for any small oracle there are local decision problems that still benefit from unique identifiers.

7.2.13. Geometry on the utility space

In [31], we study the geometrical properties of the utility space (the space of expected utilities over a finite set of options), which is commonly used to model the preferences of an agent in a situation of uncertainty. We focus on the case where the model is neutral with respect to the available options, i.e. treats them, a priori, as being symmetrical from one another. Specifically, we prove that the only Riemannian metric that respects the geometrical properties and the natural symmetries of the utility space is the round metric. This canonical metric allows to define a uniform probability over the utility space and to naturally generalize the Impartial Culture to a model with expected utilities.

7.3. Network Algorithms and Analysis

7.3.1. Information dissemination on social networks

In [17], we model an online social network as a network formation game. We study convergence of selfish dynamics and show that somewhat natural metric assumption enable fast convergence towards an equilibrium with efficient collaborative filtering of content.

7.3.2. Verification of network forwarding tables

In [25], we investigate the problem of verifying forwarding network tables. We show that it is sufficient to test few representative headers when the set of rules applied by routers is complete under intersection.

7.3.3. Refreshing old datasets in a network: LiveRank

In [18], we consider the problem of refreshing a dataset. More precisely, given a collection of nodes gathered at some time (Web pages, users from an online social network) along with some structure (hyperlinks, social relationships), we want to identify a significant fraction of the nodes that still exist at present time. The liveness of an old node can be tested through an online query at present time. We call LiveRank a ranking of the old pages so that active nodes are more likely to appear first. The quality of a LiveRank is measured by the number of queries necessary to identify a given fraction of the active nodes when using the LiveRank order. We study different scenarios from a static setting where the LiveRank is computed before any query is made, to dynamic settings where the LiveRank can be updated as queries are processed. Our results show that building on the PageRank can lead to efficient LiveRanks, for Web graphs as well as for online social networks.

7.3.4. Exploiting user movement for position detection

The major issue of indoor localization system is the trade-off between implementation cost and accuracy. A low-cost system which demands only few hardware devices could save the cost but often it turns out to be less reliable. Aiming at improving classical triangulation method that requires several reference points, we propose in [34] a new method, called Two-Step Movement (2SM), which requires only one reference point (RP) by exploiting useful information given by the position change of a mobile terminal (MT), or the user movement. This method can minimize the number of reference points required in a localization system or

navigation service and reduce system implementation cost. Analytical result shows that the user position can be thus derived and given in simple closed-form expression. Finally, simulation is conducted to demonstrate its effectiveness under noisy environment.

Then, in [35], we build on 2SM. We first improve the positioning performance through multi-sampling technique to combat measurement noise. Secondly, we propose the Generalized Two-Step Movement (G2SM) method for device-to-device (D2D) systems in which both the mobile terminal (MT) and RP can be mobile device. The mobile user's position can be derived analytically and given in simple closed-form expression. Its effectiveness in the presence of noise is shown in simulation results.

7.3.5. Fast diameter and radius computation in real-world graphs

In [5], we propose a new algorithm that computes the radius and the diameter of a weakly connected digraph $G = (V, E)$, by finding bounds through heuristics and improving them until they are validated. Although the worst-case running time is $O(|V||E|)$, we will experimentally show that it performs much better in the case of real-world networks, finding the radius and diameter values after 10–100 runs of Breadth First Search instead of $|V|$ BFS-s (independently of the value of $|V|$), and thus having running time $O(|E|)$ in practice. As far as we know, this is the first algorithm able to compute the diameter of weakly connected digraphs, apart from the naive algorithm, which runs in time $\Omega(|V||E|)$ performing a BFS from each node. In the particular cases of strongly connected directed or connected undirected graphs, we have compared our algorithm with known approaches by performing experiments on a dataset composed by several real-world networks of different kinds. These experiments show that, despite its generality, the new algorithm outperforms all previous methods, both in the radius and in the diameter computation, both in the directed and in the undirected case, both in average running time and in robustness. Finally, as an application example, we have used the new algorithm to determine the solvability over time of the “Six Degrees of Kevin Bacon” game, and of the “Six Degrees of Wikipedia” game. As a consequence, we have computed for the first time the exact value of the radius and the diameter of the whole Wikipedia digraph.

HIEPACS Project-Team

7. New Results

7.1. High-performance computing on next generation architectures

7.1.1. *Soft error sensitivity of PCG and reliability of detection mechanisms*

Soft errors can be defined as failures arising from several electricity fluctuations, cosmic particle effects on chip or any other unexpected problem while computations are in progress. If computational environment grows up to exascale, the rate of these types of error is likely to increase. These bit-flips may have a strong impact on iterative methods, that might diverge or converge to an unexpected final accuracy. Consequently, soft errors deserve to be examined in details especially in the perspective of extreme scale computing platforms. In this work, we investigate the combination of different numerical techniques to tackle the challenge of the detection. The first ingredient relies on checksum mechanisms, that are applied to secure the sparse matrix vector (SpMV) products. However, the checksum equalities are only valid in exact arithmetic while calculation are performed in finite precision. Another possibility is to monitor the residual deviation between the true and computed residual. Exploiting finite precision analysis of the round-off provides us with an upper bound on the residual norm deviation that can be used. Through intensive numerical experiments and statistical analysis we shown how round-off error analysis for the residual norm deviation can be an efficient and robust soft error detection criterion alternative to checksum approaches. This methodology has also be applied to other variants of CG, namely the pipelined and chronopolus/gear versions.

This research effort was conducted in collaboration with colleagues S. Cools and W. Vanroose from the Applied Mathematics Group of Antwerp university within the framework of the **EXA2CT** project. In this context, we also studied the impact of soft errors on a variant of the algorithm designed in their group (so-called pipelined CG). This study allowed to highlight some numerical instability in the baseline version of this variant of CG in the presence of round-off errors and we jointly proposed a correction of it that led a new both scalable and stable variant (see Section 7.2.5).

We have also designed an self-recovering CG algorithm which detects large magnituded faults with ABFT and smoothes low and average magnituded faults with deviation-based criteria.

7.1.2. *Resilience of parallel sparse hybrid solvers*

As the computational power of high performance computing (HPC) systems continues to increase by using a huge number of CPU cores or specialized processing units, extreme-scale applications are increasingly prone to faults. Consequently, the HPC community has proposed many contributions to design resilient HPC applications. These contributions may be system-oriented, theoretical or numerical. In this study we consider an actual fully-featured parallel sparse hybrid (direct/iterative) linear solver, **MaPHyS**, and we propose numerical remedies to design a resilient version of the solver. The solver being hybrid, we focus in this study on the iterative solution step, which is often the dominant step in practice. We furthermore assume that a separate mechanism ensures fault detection and that a system layer provides support for setting back the environment (processes, ...) in a running state. The present manuscript therefore focuses on (and only on) strategies for recovering lost data *after* the fault has been detected (a separate concern beyond the scope of this study), *once* the system is restored (another separate concern not studied here). The numerical remedies we propose are twofold. Whenever possible, we exploit the natural data redundancy between processes from the solver to perform exact recovery through clever copies over processes. Otherwise, data that has been lost and no longer available on any process is recovered through a so-called interpolation-restart mechanism. This mechanism is derived from our earlier studies by carefully taking into account the properties of the target hybrid solver. These numerical remedies have been implemented in the **MaPHyS** parallel solver so that we can assess their efficiency on a large number of processing units (up to 12,288 CPU cores) for solving large-scale real-life problems.

These contributions will be presented at the international conference HiPC [42].

7.1.3. Hierarchical DAG scheduling for hybrid distributed systems

Accelerator-enhanced computing platforms have drawn a lot of attention due to their massive peak computational capacity. Despite significant advances in the programming interfaces to such hybrid architectures, traditional programming paradigms struggle mapping the resulting multi-dimensional heterogeneity and the expression of algorithm parallelism, resulting in sub-optimal effective performance. Task-based programming paradigms have the capability to alleviate some of the programming challenges on distributed hybrid many-core architectures. In this work we take this concept a step further by showing that the potential of task-based programming paradigms can be greatly increased with minimal modification of the underlying runtime combined with the right algorithmic changes. We propose two novel recursive algorithmic variants for one-sided factorizations and describe the changes to the PaRSEC task-scheduling runtime to build a framework where the task granularity is dynamically adjusted to adapt the degree of available parallelism and kernel efficiency according to runtime conditions. Based on an extensive set of results we show that, with one-sided factorizations, i.e. Cholesky and QR, a carefully written algorithm, supported by an adaptive tasks-based runtime, is capable of reaching a degree of performance and scalability never achieved before in distributed hybrid environments.

These contributions will be presented at the international conference IPDPS 2015 [34] in Hyderabad.

7.1.3.1. Comparison of Static and Dynamic Resource Allocation Strategies for Matrix Multiplication

The tremendous increase in the size and heterogeneity of supercomputers makes it very difficult to predict the performance of a scheduling algorithm. In this context, relying on purely static scheduling and resource allocation strategies, that make scheduling and allocation decisions based on the dependency graph and the platform description, is expected to lead to large and unpredictable makespans whenever the behavior of the platform does not match the predictions. For this reason, the common practice in most runtime libraries is to rely on purely dynamic scheduling strategies, that make short-sighted scheduling decisions at runtime based on the estimations of the duration of the different tasks on the different available resources and on the state of the machine. In this work, we considered the special case of Matrix Multiplication, for which a number of static allocation algorithms to minimize the amount of communications have been proposed. Through a set of extensive simulations, we analyzed the behavior of static, dynamic, and hybrid strategies, and we assessed the possible benefits of introducing more static knowledge and allocation decisions in runtime libraries. These contributions have been presented at the international conference SBAC-PAD 2015.

7.1.3.2. Scheduling Trees of Malleable Tasks for Sparse Linear Algebra

Scientific workloads are often described as directed acyclic task graphs. In this paper, we focus on the multifrontal factorization of sparse matrices, whose task graph is structured as a tree of parallel tasks. Among the existing models for parallel tasks, the concept of *malleable* tasks is especially powerful as it allows each task to be processed on a time-varying number of processors. Following the model advocated by Prasanna and Musicus for matrix computations, we considered malleable tasks whose speedup is p^α , where p is the fractional share of processors on which a task executes, and α ($0 < \alpha \leq 1$) is a parameter which does not depend on the task. We first motivated the relevance of this model for our application with actual experiments on multicore platforms. Then, we studied the optimal allocation proposed by Prasanna and Musicus for makespan minimization using optimal control theory. We largely simplified their proofs by resorting only to pure scheduling arguments. Building on the insight gained thanks to these new proofs, we extended the study to distributed multicore platforms. There, a task cannot be distributed among several distributed nodes. In such a distributed setting (homogeneous or heterogeneous), we proved the NP-completeness of the corresponding scheduling problem, and proposed some approximation algorithms. We finally assessed the relevance of our approach by simulations on realistic trees. We showed that the average performance gain of our allocations with respect to existing solutions (that are thus unaware of the actual speedup functions) is up to 16% for $\alpha = 0.9$ (the value observed in the real experiments). These contributions have been presented at the international conference Europar 2015.

7.1.3.3. Task-based multifrontal QR solver for GPU-accelerated multicore architectures

Recent studies have shown the potential of task-based programming paradigms for implementing robust, scalable sparse direct solvers for modern computing platforms. Yet, designing task flows that efficiently exploit heterogeneous architectures remains highly challenging. In this work we first tackled the issue of data partitioning using a method suited for heterogeneous platforms. On the one hand, we designed task of sufficiently large granularity to obtain a good acceleration factor on GPU. On the other hand, we limited the size in order to both fit the GPU memory constraints and generate enough parallelism in the task graph. Secondly we handled the task scheduling with a strategy capable of taking into account workload and architecture heterogeneity at a reduced cost. Finally we proposed an original evaluation of the performance obtained in our solver on a test set of matrices. We showed that the proposed approach allows for processing extremely large input problems on GPU-accelerated platforms and that the overall performance is competitive with equivalent state of the art solvers designed and optimized for GPU-only use. These contributions have been presented at the international conference HiPC 2015 where they received the best paper award.

7.1.3.4. Fast and Accurate Simulation of Multithreaded Sparse Linear Algebra Solvers

The ever growing complexity and scale of parallel architectures imposes to rewrite classical monolithic HPC scientific applications and libraries as their portability and performance optimization only comes at a prohibitive cost. There is thus a recent and general trend in using instead a modular approach where numerical algorithms are written at a high level independently of the hardware architecture as Directed Acyclic Graphs (DAG) of tasks. A task-based runtime system then dynamically schedules the resulting DAG on the different computing resources, automatically taking care of data movement and taking into account the possible speed heterogeneity and variability. Evaluating the performance of such complex and dynamic systems is extremely challenging especially for irregular codes. In this work, we explained how we crafted a faithful simulation, both in terms of performance and memory usage, of the behavior of `qr_mumps`, a fully-featured sparse linear algebra library, on multi-core architectures. In our approach, the target high-end machines are calibrated only once to derive sound performance models. These models can then be used at will to quickly predict and study in a reproducible way the performance of such irregular and resource-demanding applications using solely a commodity laptop. These contributions have been presented at the international conference ICPADS 2015.

7.2. High performance solvers for large linear algebra problems

7.2.1. Divide and conquer symmetric tridiagonal eigensolver for multicore architectures

Computing eigenpairs of a symmetric matrix is a problem arising in many industrial applications, including quantum physics and finite-elements computation for automobiles. A classical approach is to reduce the matrix to tridiagonal form before computing eigenpairs of the tridiagonal matrix. Then, a back-transformation allows one to obtain the final solution. Parallelism issues of the reduction stage have already been tackled in different shared-memory libraries. In this work, we focus on solving the tridiagonal eigenproblem, and we describe a novel implementation of the Divide and Conquer algorithm. The algorithm is expressed as a sequential task-flow, scheduled in an out-of-order fashion by a dynamic runtime which allows the programmer to play with tasks granularity. The resulting implementation is between two and five times faster than the equivalent routine from the INTEL MKL library, and outperforms the best MRRR implementation for many matrices. These contributions have been presented at the international conference IPDPS 2015 [32] in Hyderabad.

7.2.2. Blocking strategy optimizations for sparse direct linear solver on heterogeneous architectures

Solving sparse linear systems is a problem that arises in many scientific applications, and sparse direct solvers are a time consuming and key kernel to those applications or more advanced solvers such as hybrid direct-iterative solvers. That is why optimizing their performance on modern architectures is a crucial problem. The preprocessing steps of sparse direct solvers: ordering and symbolic factorization, are two major steps that lead to a reduced amount of computation and memory, and to a better task granularity to reach a good level of performance when using BLAS kernels. With the advent of GPUs, the granularity of the symbolic factorization

became more important than ever. In this work, we present a reordering strategy that increases the block granularity. This strategy relies on the symbolic factorization to refine the ordering produced by tools such as METIS or **Scotch**, and does not impact the number of operations required to solve the problem. We integrated this algorithm in the **PaStiX** solver and show a reduction of the number of off-diagonal blocks by two to three on a large spectrum of matrices. This improvement leads to an efficiency on GPUs raised by up to 40%. These contributions have been presented at the Sparse Days [51] in Saint-Girons.

7.2.3. *On the use of \mathcal{H} -Matrix Arithmetic in PaStiX: a Preliminary Study*

The objective is to investigate innovative lowrank approximations based on \mathcal{H} -matrix variants for direct solver and Schur complements. The intent is to improve scalability of those components involved in preconditioners and hybrid solvers by reducing the computational and memory costs of the dense calculation. The quality of hybrid ordering algorithms combining topdown (such as nested dissection) and bottomup (such as minimum degree) ordering techniques in the context of sparse linear solvers will be investigated.

In this work, we describe a preliminary fast direct solver using HODLR library to compress large blocks appearing in the symbolic structure of the **PaStiX** sparse direct solver. We present our general strategy before analyzing the practical gains in terms of memory and floating point operations with respect to a theoretical study of the problem. Finally, we discuss ways to enhance the overall performance of the solver.

Some contributions have already been presented at the Workshop on Fast Solvers [52] in Toulouse. This work is a joint effort between Professor Darve's group at Stanford and the Inria HiePACS team within **FASTLA**.

7.2.4. *Data sparse techniques for parallel hybrid solvers*

In this work we describe how data sparse techniques exploiting \mathcal{H} -matrix calculations can be implemented in a parallel hybrid sparse linear solver based on an algebraic non overlapping domain decomposition approach.

Various graph-based clustering techniques to approximate the local Schur complements are investigated, with the aim of optimally complying with the interface structure of the local interfaces of the subdomains. We consider strong-hierarchical (sH) matrix arithmetic as efficient means for obtaining low rank approximations in terms of workload distribution as well as memory consumption. We also show how sH-arithmetic can be utilized to form an effective global preconditioner for the iterative phase of the hybrid solver. Numerical and parallel experiments are presented to evaluate the advantages and drawbacks of the different variants.

This work is a joint effort between Professor Darve's group at Stanford and the Inria HiePACS team within **FASTLA**. Some intermediate progresses have already been presented [38], [37]

7.2.5. *Analysis of the rounding error accumulation in Conjugate Gradient to improve the maximal attainable accuracy of pipelined CG*

Pipelined Krylov solvers typically offer better scalability in the strong scaling limit compared to standard Krylov methods. The synchronization bottleneck is mitigated by overlapping time-consuming global communications with useful computations in the algorithm. However, to achieve this communication hiding strategy, pipelined methods feature multiple recurrence relations on additional auxiliary variables to update the guess for the solution. This paper aims to study the influence of rounding errors on the convergence of the pipelined Conjugate Gradient method. It is analyzed why rounding effects have a significantly larger impact on the maximal attainable accuracy of the pipelined CG algorithm compared to the traditional CG method. Based on a rounding error model, we then propose an automated residual replacement strategy to reduce the effect of rounding errors on the final iterative solution. The resulting pipelined CG method with residual replacement improves the maximal attainable accuracy of pipelined CG while maintaining its efficient parallel performance.

This research effort was conducted in collaboration with colleagues S. Cools and W. Vanroose from the Applied Mathematics Group of Antwerp university within the framework of the **EXA2CT** project.

7.3. High performance Fast Multipole Method for N-body problems

7.3.1. Task-based Fast Multipole Method

Last year we have worked primarily on developing an efficient fast multipole method for heterogeneous architecture. Some of the accomplishments for this year include:

1. We have finalized the Uniform FMM (ufmm) based on polynomial interpolations combined with a hierarchical (data sparse) representation of a kernel matrix. The algorithm is close to the Black Box FMM by Fong and Darve developed with Chebyshev polynomials, however it uses an interpolation scheme based on an equispaced grid, which allows the use of FFT and consequently reduce both running time and memory footprint but has implications on accuracy and stability. The theory behind the Uniform FMM kernel is explained in a research report [63] along with numerical benchmarks on artificial test cases and presented in [44]. This new kernel was extended to be used for dislocation kernel.
2. Concerning the Group-Tree approach, we have shown in past studies its advantages of the task-based FMM and how the group-tree is well suited for runtime systems. In fact, it improves the locality, but it also reduces the number of dependencies which is an important asset to decrease the runtime overhead. These prospective task-based FMM can solve problems on heterogeneous architecture as presented in [36]. Therefore, we have continued this work and created a robust group-tree that has been included in ScalFMM and which is now available to the community. This data structure is generic and can be used with the different ScalFMM kernels. Moreover, we have extended our work and implemented a distributed task-based FMM above StarPU. The description of the data structure and some experimental studies will be presented in February 2016 during PhD defense of B. Bérenger.
3. With the advent of complex modern architectures, the low-level paradigms long sufficient to build high performance computing (HPC) numerical codes have met their limits. Achieving efficiency, ensuring portability, while preserving programming tractability on such hardware prompted the HPC community to design new, higher level paradigms. Indeed, several robust runtime systems proposed recently have shown the benefit of task-based parallelism models in terms of performance portability on complex platforms, on top of which full-featured numerical libraries have been ported successfully. However, the common weakness of these projects is to deeply tie applications to specific expert-only runtime system APIs. The OPENMP specification, which aims at providing a common parallel programming means for shared-memory platforms, appears a good candidate to address this issue thanks to the latest task-based constructs introduced as part of its revision 4.0. The goal of this joint work with STORM team is to assess the effectiveness and limits of this support for designing a high-performance numerical library like ScalFMM library, which implements state-of-the-art fast multipole methods (FMM) algorithms and that we have considerably re-designed with respect to the most advanced features provided by OPENMP 4.0. We show that OPENMP 4.0 allows for significant performance improvements over previous OPENMP revisions on recent multicore processors. We furthermore propose extensions to the OPENMP 4.0 standard and show how they could enhance FMM performance. To assess our statement, we have implemented this support within the KLANG-OMP source-to-source compiler that translates OPENMP directives into calls to the StarPU task-based runtime system. This study shows that we can take advantage of the advanced capabilities of a fully-featured runtime system without resorting to a specific, native runtime port, hence bridging the gap between the OPENMP standard and the very high performance that was so far reserved to expert-only runtime system APIs.

7.3.2. Time-domain boundary element method

The Time-domain Boundary Element Method (TD-BEM) has not been widely studied but represents an interesting alternative to its frequency counterpart. Usually based on inefficient Sparse Matrix Vector-product (SpMV), we investigate other approaches in order to increase the sequential flop-rate.

The TD-BEM formulation we is naturally expressed using sparse-matrix vector product (SpMV). We describe how the Flop-rate can be improved using a so-called multi-vectors/vector product, and we provide an efficient implementation of this operation using vectorization. We have extended our TD-BEM solver to support NVidia GPUs, and we have looked at different blocking schemes and their respective implementations. We have created a new blocking storage which matches our operators and allows to obtain a high Flop-rate. In addition, we provide a balancing heuristic to divide the work between the CPUs and the GPUs dynamically. The results have been published in [20], and our solver is now able to work on distributed heterogeneous nodes.

Our TD-BEM solver is efficient, but it still has a quadratic complexity which might become a problem for large problems. This high complexity motivates the study of an FMM based TD-BEM solver with the objective of being more competitive as the problem size increases. Therefore, we have implemented an FMM-based solver but while the complexity should be lower than the matrix approach, it remains unclear from which problem size. Moreover, we show in [PhD defense of B. Béranger] different results and point-out that the memory cost is much more expensive for the FMM approach compare to the matrix one. The method has been discussed in [43] among other ScalFMM applications.

All the implementations should be in high quality in the Software Engineering sense since the resulting library is going to be used by industrial applications.

This work is developed in the framework of Béranger Bramas's PhD and contributes to the EADS-ASTRIUM, Inria, Conseil Régional initiative.

7.3.3. Randomized algorithms for covariance matrices

7.3.3.1. Covariance kernel matrices

Random projection based Low Rank Approximation (LRA) algorithms such as the randomized SVD produce approximate matrix factorizations in quadratic instead of cubic time in N (N being the matrix size). This complexity can be further improved if fast matrix multiplication is available. A paper explaining our recent advances in fast randomized LRA of covariance kernel matrices using FMM is available as a research report [63] and presented in [44]. In particular, the fast multipole acceleration of the randomized SVD allowed for generating Gaussian random fields on arbitrary grids in linear running time and memory requirements. The code is available in the open source C++ project FMR: <https://gforge.inria.fr/projects/fmr>, it relies heavily on the ScalFMM library for data structures and fast matrix multiplication.

7.3.3.2. New applications: Data Assimilation and Taxonomy

Many applications like data assimilation (e.g. Kalman Filtering or variational approaches) or biology (e.g. taxonomy) involve covariance matrices that are only known in algebraic form, as opposed to kernel matrices that can be explicitly build given a kernel function. In a joint project (called FastMDS) with Alain Franc (INRA, Inria PLEIADE) addressing fast methods for the classification of biological species (taxonomy) our randomized SVD algorithm was used in order to accelerate a MultiDimensionalScaling (MDS) algorithm. The MDS is a widely used method in machine learning and data analysis that aim at visualizing the information contained in a distance matrix. Our MDS algorithm is applied to DNA sequences coming from various sources (e.g. Leman's lake), it consists in forming an euclidian image of the sample by taking the square root of a covariance matrix computed from the distance matrix. The randomized SVD approach lead to promising results, since it allowed to treat up to 100.000 samples in a few seconds. Since the covariance matrix still needs to be loaded in memory, storage might become problematic for larger samples. Therefore we are now considering matrix-free methods in order to decrease the memory requirements but also hierarchical algorithms in order to compute the MDS in near-linear time. The following methods are currently under investigation:

- Random column selection based LRA methods such as the Nystrom method or blocked variant of the Nystrom method (BBF, see Wang, Darve, Mahoney).
- Random projection based LRA powered by general H2-methods.

All these techniques are considered since they apply well, when the relevant information is spread uniformly among the data, just like in our data sets.

7.4. Efficient algorithmic for load balancing and code coupling in complex simulations

7.4.1. Dynamic load balancing for massively parallel coupled codes

In the field of scientific computing, load balancing is a major issue that determines the performance of parallel applications. Nowadays, simulations of real-life problems are becoming more and more complex, involving numerous coupled codes, representing different models. In this context, reaching high performance can be a great challenge. In the PhD of Maria Predari (started in October 2013), we develop new graph partitioning techniques, called co-partitioning, that address the problem of load balancing for two coupled codes: the key idea is to perform a *coupling-aware* partitioning, instead of partitioning these codes independently, as it is usually done. However, our co-partitioning technique requires to use graph partitioning with *fixed vertices*, that raises serious issues with state-of-the-art software, that are classically based on the well-known recursive bisection paradigm (RB). Indeed, the RB method often fails to produce partitions of good quality. To overcome this issue, we propose a *new* direct *k*-way greedy graph growing algorithm, called KGGGP, that overcomes this issue and succeeds to produce partition with better quality than RB while respecting the constraint of fixed vertices. Experimental results compare KGGGP against state-of-the-art methods for graphs available from the popular DIMACS'10 collection. This work will be presented in the 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP 2016).

7.5. Application Domains

7.5.1. Material physics

7.5.1.1. Molecular Vibrational Spectroscopy

Quantum chemistry eigenvalue problem is a big challenge in recent research. Here we are interested in solving eigenvalue problems coming from the molecular vibrational analysis. These problems are challenging because the size of the vibrational Hamiltonian matrix to be diagonalized is exponentially increasing with the size of the molecule we are studying. So, for molecules bigger than 10 atoms the actual existent algorithms suffer from a curse of dimensionality or computational time. We propose a new variational algorithm (namely residue-based adaptive vibrational configuration interaction) intended for the resolution of the vibrational Schrödinger equation. The main advantage of this approach is to efficiently reduce the dimension of the active space generated into the configuration interaction (CI) process. This adaptive algorithm is developed with the use of three correlated conditions i.e. a suitable starting space ; a criterion for convergence, and a procedure to expand the approximate space. The speed of the algorithm was increased with the use of a posteriori error estimator (residue) to select the most relevant direction to increase the space. Two examples have been selected for benchmark. In the case of Formalde hydemolecule (H_2CO) with a dimension space of 6, we mainly study the performance of RA-VCI algorithm: comparison with the variation-perturbation method, choice of the initial space, residual contributions. For Acetonitrile molecule (CH_3CN) with dimension space of 12 the active space computed by our algorithm is divided by 20 compared to the computations done by Avila et. al using the same potential energy surface. This work was presented in [54], [53].

7.5.1.2. Dislocations

7.5.1.2.1. Direct evaluation of the anisotropic elastic force field

The anisotropic elastic force field created by dislocations is not explicitly given, in fact it is only known in integral form using Green's or Stroh's formalism. The approach considered in OptiDis is based on Stroh's formalism, i.e. we compute the stress field using tensorial angular functions known as Stroh matrices. A benefit of using Stroh's formalism is that it only requires the evaluation of a single line integral for the force field and no integration for the stress field, while Green's formalism involve double and single line integral respectively. The evaluation of Stroh matrices in arbitrary directions is not affordable, therefore spherical harmonic expansions were considered in order to approximate the stress field efficiently. Until now the integration of the stress field on target dislocations was performed numerically using simple quadratures,

although the quadrature size required to evaluate the force field at a given precision may explode as segments get closer and computation may become untractable. In order to avoid this behaviour, we developed semi-analytical expressions of the force field based on the analytic integration of the expansions of the stress field (in spherical harmonics). This new method is an adaptation of Aubry et al. approach to Stroh's formalism, in the sense that it also provides optimized recursive formulae to efficiently evaluate these semi-analytic expressions. Numerous verifications and further improvements of the expressions are required before implementing it inside OptiDis.

7.5.1.2.2. Parallel dislocation dynamics simulation

We have focused on the improvements of our hybrid MPI-OpenMP parallelism of the OptiDis code. More precisely, we have continued the development of parallel algorithm to add/remove element in the cache-conscious data structure. This data structured combined with an octree manages efficiently large set of data (segments and nodes) during all the steps of the algorithm. Moreover, we have tuned and improved our hybrid MPI-OpenMP parallelism to run simulations with large number of radiation induced defects forming our dislocation network. To obtain a good scalability, we have introduced a better load balancing at thread level as well as process level. By combining efficient data structure and hybrid parallelism we obtained a speedup of 112 on 160 cores for a simulation of half a million of segments.

All this work was developped in the Phd of A. Etchevery.

7.5.2. Co-design for scalable numerical algorithms in scientific applications

7.5.2.1. MHD instabilities edge localized modes

The last contribution of Xavier Lacoste's thesis deals with the integration of our work in **JOEK**, a production controlled plasma fusion simulation code from CEA Cadarache. We described a generic finite element oriented distributed matrix assembly and solver management API. The goal of this API is to optimize and simplify the construction of a distributed matrix which, given as an input to **PaStiX**, can improve the memory scaling of the application. Experiments exhibit that using this API we could reduce the memory consumption by moving to a distributed matrix input and improve the performance of the factorized matrix assembly by reducing the volume of communication. All this study is related to **PaStiX** integration inside **JOEK** but the same API could be used to produce a distributed assembly for another solver or/and another finite elements based simulation code.

7.5.2.2. Turbulence of plasma particules inside a tokamak

Concerning the **GSELA** global non-linear electrostatic code, the efforts during the period have concentrated on predicting memory requirement and on the gyroaverage operator.

The Gysela program uses a mesh of 5 dimensions of the phase space (3 dimensions in configuration space and 2 dimensions in velocity space). On the large cases, the memory consumption already reaches the limit of the available memory on the supercomputers used in production (Tier-1 and Tier-0 typically). Furthermore, to implement the next features of Gysela (e.g. adding kinetic electrons in addition to ions), the needs of memory will dramatically increase, the main unknown will represents hundreds of TB. In this context, two tools were created to analyze and decrease the memory consumption. The first one is a tool that plots the memory consumption of the code during a run. This tool helps the developer to localize where the memory peak is located. The second tool is a prediction tool to compute the peak memory in offline mode (for production use mainly). A post processing stage combined with some specific traces generated on purpose during runtime allow the analysis of the memory consumption. Low-level primitives are called to generate these traces and to model memory consumption : they are included in the libMTM library (Modeling and Tracing Memory). Thanks to this work on memory consumption modeling, we have decreased the memory peak of the Gysela code up to 50 % on a large case using 32,768 cores and memory scalability improvement has been shown using these tools up to 65k cores.

The main unknown of the Gysela is a distribution function that represents either the density of the guiding centers, either the density of the particles in a tokamak (depending of the location in the code). The switch between these two representations is done thanks to the gyroaverage operator. In the actual version of Gysela, the computation of this operator is achieved thanks to the so-called Padé approximation. In order to improve the precision of the gyroaveraging, a new implementation based on interpolation methods has been done (mainly by researchers from the Inria Tonus project-team and IPP Garching). We have performed the integration of this new implementation in Gysela and also some parallel benchmarks. However, the new gyroaverage operator is approximatively 10 times slower than the original one. Investigations and optimizations on this operator are still a work in progress.

This work has been carried on in the framework of Fabien Rozar's PhD in collaboration with CEA Cadarache (defended in November 2015). A new PhD (Nicolas Bouzat) has started in October 2015 and the scientific objectives of this work will be first to consolidate the parallel version of the gyroaverage operator, in particular by designing a complete MPI+OpenMP parallel version, and then to design new numerical methods for the gyroaverage, source and collision operators to deal with new physics in Gysela. The objective is to tackle kinetic electron configurations for more realistic simulations.

7.5.2.3. SN Cartesian solver for nuclear core simulation

High-fidelity nuclear power plant core simulations require solving the Boltzmann transport equation. In discrete ordinate methods, the most computationally demanding operation of this equation is the sweep operation. Considering the evolution of computer architectures, we propose in this work, as a first step toward heterogeneous distributed architectures, a hybrid parallel implementation of the sweep operation on top of the generic task-based runtime system: **PaRSEC**. Such an implementation targets three nested levels of parallelism: message passing, multi-threading, and vectorization. A theoretical performance model was designed to validate the approach and help the tuning of the multiple parameters involved in such an approach. The proposed parallel implementation of the Sweep achieves a sustained performance of 6.1 Tflop/s, corresponding to 33.9% of the peak performance of the targeted supercomputer. This implementation compares favorably with state-of-art solvers such as PARTISN; and it can therefore serve as a building block for a massively parallel version of the neutron transport solver DOMINO developed at EDF.

The main contribution has been presented at the international conference IPDPS 2015 [31] in Hyderabad.

7.5.2.4. 3D aerodynamics for unsteady problems with moving bodies

In the first part of our research work concerning the parallel aerodynamic code FLUSEPA, a first OpenMP-MPI version based on the previous one has been developed. By using an hybrid approach based on a domain decomposition, we achieved a faster version of the code and the temporal adaptive method used without bodies in relative motion has been tested successfully for real complex 3D-cases using up to 400 cores. Moreover, an asynchronous strategy for computing bodies in relative motion and mesh intersections has been developed and has been used for actual 3D-cases. A journal article (for JCP) to sum-up this part of the work is under redaction and a presentation at ISC at the "2nd International Workshop on High Performance Computing Simulation in Energy/Transport Domains" on July 2015 is scheduled.

This intermediate version exhibited synchronization problems for the aerodynamic solver due to the time integration used by the code. To tackle this issue, a task-based version over the runtime system **StarPU** is currently under development and evaluation. This year was mainly devoted to the realisation of this version. Task generation function have been designed in order to maximize asynchronism in execution. Those functions respect the data pattern access of the code and led to the refactorization of the actual kernels. A task-based version is now available for the aerodynamic solver and is available for both shared and distributed memory. This work has been presented as a poster during the SIAM CSE' 15 conference and at the Parallel CFD' 15 and HPCSET' 15 conferences.

The next steps will be to validate the correction of this task-based version and to work on the performance of this new version on actual cases. Later, the task description should be extended to the motion and intersection operations.

This work is carried on in the framework of Jean-Marie Cousteyen's PhD in collaboration with Airbus Defence and Space.

7.5.2.5. Spectral recycling strategies for the solution of nonlinear eigenproblems in thermoacoustics

In this work we consider the numerical solution of large nonlinear eigenvalue problems that arise in thermoacoustic simulations involved in the stability analysis of large combustion devices. We briefly introduce the physical modeling that leads to a nonlinear eigenvalue problem that is solved using a nonlinear fixed point iteration scheme. Each step of this nonlinear method requires the solution of a complex non-Hermitian linear eigenvalue problem. We review a set of state of the art eigensolvers and discuss strategies to recycle spectral information from one nonlinear step to the next. More precisely, we consider the Jacobi-Davidson algorithm, the Implicitly Restarted Arnoldi method, the Krylov-Schur solver and its block-variant, as well as the subspace iteration method with Chebyshev acceleration. On a small test example we study the relevance of the different approaches and illustrate on a large industrial test case the performance of the parallel solvers best suited to recycle spectral information for large scale thermoacoustic stability analysis.

The results of this work conducted in collaboration with S. Moreau (Sherbrooke University) and Y. Saad (University of Minnesota Twin-cities) are detailed in [22]

7.5.2.6. A conservative 2-D advection model towards large-scale parallel calculation

To exploit the possibilities of parallel computers, we designed a large-scale bidimensional atmospheric advection model named Pangolin. As the basis for a future chemistry-transport model, a finite-volume approach for advection was chosen to ensure mass preservation and to ease parallelization. To overcome the pole restriction on time steps for a regular latitude-longitude grid, Pangolin uses a quasi-area-preserving reduced latitude-longitude grid. The features of the regular grid are exploited to reduce the memory footprint and enable effective parallel performances. In addition, a custom domain decomposition algorithm is presented. To assess the validity of the advection scheme, its results are compared with state-of-the-art models on algebraic test cases. Finally, parallel performances are shown in terms of strong scaling and confirm the efficient scalability up to a few hundred cores

The results of this work are detailed in [21].

INDES Project-Team

6. New Results

6.1. Web programming

Participants: Yoann Couillec, Vincent Prunet, Manuel Serrano [correspondant].

6.1.1. Hop.js

Multitier programming languages unify within a single formalism and a single execution environment the programming of the different tiers of distributed applications. On the Web, this programming paradigm unifies the client tier, the server tier, and, when one is used, the database tier. This homogenization offers several advantages over traditional Web programming that rely on different languages and different environments for the two or three tiers of the Web application: programmers have only one language to learn, maintenance and evolution are simplified by the use of a single formalism, global static analyses are doable as a single semantics is involved, debugging and other runtime tools are more powerful as they access global informations about the execution.

The three first multitier platforms for the Web all appeared in 2006: GWT (a.k.a., Google Web Toolkit), Links, and Hop [6], [5]. Each relied on a different programming model and languages. GWT maps the Java programming model on the Web, as it allows, Java/Swing likes programs to be compiled and executed on the Web; Links is functional language with experimental features such as the storing of the whole execution context on the client; Hop is based on the Scheme programming language. These three pioneers have open the path for the other multitier languages such as, Ocsigen for Ocaml, UrWeb, js-scala, etc.

In spite of their interesting properties, multitier languages have not become that popular on the Web. Today, only GWT is widely used in industrial applications but arguably GWT is not a fully multitier language as developing applications with GWT requires explicit JavaScript and HTML programming. This lack of popularity of other systems is likely due to their core based languages than to the programming model itself.

JavaScript is the *de facto* standard on the Web. Since the mid 90's, it is the language of the client-side programming and more recently, with systems like Node.js, it is also a viable solution for the server-side programming. As we are convinced by the virtues of multitier programming we have started a new project consisting of enabling multitier programming JavaScript. We have created a new language called HopScript, which is a minimalist extension of JavaScript for multitier programming, and we have implemented a brand new runtime environment called Hop.js. This environment contains a builtin Web server, on-the-fly HopScript compilers, and many runtime libraries.

HopScript is a super set of JavaScript, *i.e.*, all JavaScript programs are legal HopScript programs. Hop.js is a compliant JavaScript execution environment as it succeeds at 99% of the Ecma 262 tests suite. The Hop.js environment also aims at Node.js compatibility. In its current version it supports about 70% of the Node.js runtime environment. In particular, it fully supports the Node.js modules, which lets Hop programs reuse existing Node.js modules as is.

After a full year of active development to enhance JavaScript and Node.js compatibility, to incorporate features of JavaScript 1.6, and to design new language constructs for machine-to-machine communication, we are now ready to release Hop.js. This will appear at the beginning of 2016.

6.1.2. Data source

During the past few years the volume of accumulated data has increased dramatically. New kinds of data stores have emerged as NoSQL family stores. Many modern applications now collect, analyze, and produce data from several heterogeneous sources. However implementing such applications is still difficult because of lack of appropriate tools and formalisms. We propose a solution to this problem in the context of the JavaScript

programming language by extending array comprehensions. Our extension allows programmers to query data from usual stores, such as SQL databases, NoSQL databases, Semantic Web data repositories, Web pages, or even custom user defined data structures. The extension has been implemented in the Hop.js system. It has been described in the paper [10], which has been presented at the ACM DBPL'15 conference.

6.2. Distributed programming

Participants: Gérard Boudol, Johan Grande, Manuel Serrano [correspondant].

Shared-memory concurrency is a classic concurrency model which, among other things, makes it possible to take advantage of multicore processors that are now widespread in personal computers. Concurrent programs are prone to deadlocks which are notoriously hard to predict and debug. Programs using mutexes, a very popular synchronization mechanism, are no exception.

We have studied deadlock avoidance methods with the aim of making programming with mutexes easier. We first studied a method that uses a static analysis by means of a type and effect system, then a variation on this method in a dynamically typed language.

We developed more the second method. It mixes deadlock prevention and avoidance to provide an easy-to-use and expressive deadlock-free locking function. We implemented it as a Hop library. This led us to develop a starvation-free algorithm to simultaneously acquire an arbitrary number of mutexes, and to identify the concept of asymptotic deadlock. While doing so, we also developed an optimization of exceptions (finally blocks).

Our performance tests seem to show that using our library has negligible impact on the performance of real-life applications. Most of our work could be applied to other structured programming languages such as Java.

This work has been presented at the 17th International Symposium on Principles and Practice of Declarative Programming (PPDP'15) [13]. More details can be found in Grande's PhD thesis [8].

6.3. Types

Participants: Ilaria Castellani, Bernard Serpette.

6.3.1. Behavioural Types

The survey paper <https://hal.inria.fr/hal-01213201> presents a state-of-the-art of a recent trend of research on the use of behavioural types for specifying and analysing security properties of communication-centred systems. It is essentially an outcome of the working group on security of the BETTY COST Action, and it offers a unified overview of various proposals that have been put forward in the last few years, both within the BETTY community and outside it, to combine security analysis with behavioural types.

6.3.2. Abstract Rewriting Systems

We have formalised, with the Coq system, the beginning of Paul-André Melliès's thesis concerning abstract rewriting systems. Behind the interest of studying rewriting systems, which are the roots of all small step semantics of programming languages, this particular formalisation was attractive since it gives a concrete example where we have to manage dependant types.

This was done in collaboration with Eduardo Bonelli and Pablo Barenbaum of University of Quilmes, Argentina. The specification and the proofs of this work take 2200 lines of Coq.

6.4. Security

Participants: Ilaria Castellani, Francis Doliere Some, Nataliia Bielova, Bernard Serpette, Tamara Rezk [correspondant].

6.4.1. Hybrid Typing of Secure Information Flow in a JavaScript-like Language

We propose a novel type system for securing information flow in a core of JavaScript. This core takes into account the defining features of the language, such as prototypical inheritance, extensible objects, and constructs that check the existence of object properties. We design a hybrid version of the proposed type system. This version infers a set of assertions under which a program can be securely accepted and instruments it so as to dynamically check whether these assertions hold. By deferring rejection to runtime, the hybrid version can typecheck secure programs that purely static type systems cannot accept.

This work has been published at the 10th International Symposium on Trustworthy Global Computing [11].

6.4.2. Modular Monitor Extensions for Information Flow Security in JavaScript

Client-side JavaScript programs often interact with the web page into which they are included, as well as with the browser itself, through APIs such as the DOM API, the XMLHttpRequest API, and the W3C Geolocation API. Precise reasoning about JavaScript security must therefore take API invocation into account. However, the continuous emergence of new APIs, and the heterogeneity of their forms and features, renders API behavior a moving target that is particularly hard to capture. To tackle this problem, we propose a methodology for modularly extending sound JavaScript information flow monitors with a generic API. Hence, to verify whether an extended monitor complies with the proposed noninterference property, our methodology requires only to prove that the API satisfies a predefined set of conditions. In order to illustrate the practicality of our methodology, we show how an information flow monitor-inlining compiler can take into account the invocation of arbitrary APIs, without changing the code or the proofs of the original compiler. We provide an implementation of such a compiler with an extension for handling a fragment of the DOM Core Level 1 API. Furthermore, our implementation supports the addition of monitor extensions for new APIs at runtime. This work has been published at the 10th International Symposium on Trustworthy Global Computing [12].

6.4.3. Relaxed Noninterference

We have begun a study concerning the use of gradual typing for down casting or declassification for information flow. The particularity of this work is to use a finite state machine to gradually accept the down casting process.

This work is done with Éric Tanter of University of Santiago de Chile, in the context of the project Conicyt Redes CEV Challenges on Electronic Voting.

6.4.4. Hybrid Monitoring of Attacker knowledge

Enforcement of non-interference requires to prove that an attacker's knowledge about the initial state remains the same after observing a program's public output. We define a powerful hybrid monitoring mechanism which evaluates dynamically the knowledge that is contained in program variables. To get a precise estimate of the knowledge, the monitor statically analyses non-executed branches. We show that our knowledge-based approach can be combined with existing dynamic monitors for non-interference. A distinguishing feature of such a combination is that the combined monitor is provably more powerful than each mechanism taken separately. We demonstrate this by proposing a knowledge-enhanced version of a dynamic monitor based on the no-sensitive-upgrade principle. We show how to use the knowledge computed by our hybrid monitor to quantify information leakage associated to the program output. The monitor and its static analysis has been formalized and proved correct within the Coq proof assistant.

6.4.5. A Taxonomy of Information Flow Monitors

We propose a rigorous comparison of information flow monitors with respect to two dimensions: soundness and transparency.

For soundness, we notice that the standard information flow security definition called *Termination-Insensitive Non-interference (TINI)* allows the presence of termination channels, however it does not describe whether the termination channel was present in the original program, or it was added by a monitor. We propose a stronger notion of noninterference, that we call *Termination-Aware Non-interference (TANI)*, that captures this fact, and thus allows us to better evaluate the security guarantees of different monitors. We further investigate TANI, and state its formal relations to other soundness guarantees of information flow monitors. For transparency, we identify different notions from the literature that aim at comparing the behaviour of monitors. We notice that one common notion used in the literature is not adequate since it identifies as better a monitor that accepts insecure executions, and hence may augment the knowledge of the attacker. To discriminate between monitors' behaviours on secure and insecure executions, we factorized two notions that we call true and false transparency. These notions allow us to compare monitors that were deemed to be incomparable in the past.

We analyse five widely explored information flow monitors: no-sensitive- upgrade (NSU), permissive-upgrade (PU), hybrid monitoring (HM), secure multi-execution (SME), and multiple facets (MF).

This work has been accepted for publication in the International Conference on Principles of Security and Trust (POST 2016).

6.4.6. A Study of JavaScript constructs used in Top Alexa Sites

Several works on JavaScript analysis have shown that including remote scripts can introduce severe security implications in the behavior of the whole web application. To deal with different kinds of attacks, a number of research groups are developing automatic tools to analyze JavaScript programs. However, most of these works rely on one assumption: the scripts are written in a subset of JavaScript language meaning that only certain constructs are used (that are easier to analyse automatically) and others are omitted (for example, `eval` is impossible to analyze statically). The goal of the internship was to account for the use of each JavaScript construct in real world programs. To achieve that, we first did a large-scale crawl of the top 10,000 Alexa sites, collecting both inlined scripts and remote scripts. Second, we established the popularity of remote scripts. Next, we accounted for the occurrence of JavaScript constructs in the collected programs. Finally, we use the occurrence of different constructs as basis to propose a subset of JavaScript language, which covers most of JavaScript programs found in the wild. One can rely on this evidence-based subset of JavaScript in future works on that language.

INFINE Team

6. New Results

6.1. Online Social Networks (OSN)

Community detection; bandit algorithms; privacy preservation; reward mechanisms

6.1.1. *Community detection*

Participants: Laurent Massoulié, Marc Lelarge, Charles Bordenave.

We have progressed in the design of spectral methods for community detection and in the corresponding analysis, in particular by proving the so-called spectral redemption conjecture. This has been published in IEEE FOCS'15. The abstract of the paper is as follows. A non-backtracking walk on a graph is a directed path such that no edge is the inverse of its preceding edge. The nonbacktracking matrix of a graph is indexed by its directed edges and can be used to count non-backtracking walks of a given length. It has been used recently in the context of community detection and has appeared previously in connection with the Ihara zeta function and in some generalizations of Ramanujan graphs. In this work, we study the largest eigenvalues of the non-backtracking matrix of the Erdős-Rényi random graph and of the Stochastic Block Model in the regime where the number ℓ of edges is proportional to the number of vertices. Our results confirm the “spectral redemption conjecture” that community detection can be made on the basis of the leading eigenvectors above the feasibility threshold.

6.1.2. *Bandit algorithms for active learning of content type at low spam cost*

Participants: Laurent Massoulié, Mesrob Ohanessian, Alexandre Proutière.

Progress on “bandit algorithms” for targeted news dissemination. We developed a framework in which to cast the problem, and the so-called “greedy Bayes” algorithm to determine which user to expose to a given content. We proved corresponding optimality properties, and observed that “greedy Bayes” beats the so-called Thompson sampling approach, that is the state-of-the-art method in bandit problems. This work was published at ACM Sigmetrics'15.

6.1.3. *Clustering and Inference From Pairwise Comparisons*

Participants: Rui Wu, Jiaming Xu, Srikant Rayadurgam, Marc Lelarge, Laurent Massoulié, Bruce Hajek.

In a short publication at ACM Sigmetrics'15, we do the following. Given a set of pairwise comparisons, the classical ranking problem computes a single ranking that best represents the preferences of all users. In this paper, we study the problem of inferring individual preferences, arising in the context of making personalized recommendations. In particular, we assume users form clusters; users of the same cluster provide similar pairwise comparisons for the items according to the Bradley-Terry model. We propose an efficient algorithm to estimate the preference for each user: first, compute the net-win vector for each user using the comparisons; second, cluster the users based on the net-win vectors; third, estimate a single preference for each cluster separately. We show that the net-win vectors are much less noisy than the high dimensional vectors of pairwise comparisons, therefore our algorithm can cluster the users reliably. Moreover, we show that, when a cluster is only approximately correct, the maximum likelihood estimation for the Bradley-Terry model is still close to the true preference.

6.2. Spontaneous Wireless Networks and Internet of Things

internet of things; wireless sensor networks; dissemination; resource management

6.2.1. *Platform Design for the Internet of Things*

Participants: Emmanuel Baccelli, Cedric Adjih, Oliver Hahm, Matthias Waehlich, Thomas Schmidt, Hauke Petersen.

Within this activity, we have further developed the platforms we champion for the Internet of Things: the open source operating system RIOT and open-access IoT-lab testbeds. RIOT now aggregates open source contributions from 120+ people (and counting) from all over the world, coming both from academia and from industry, and received financial backing from top companies including Cisco and Google in 2015. Revisiting concepts from the early Internet, we have designed and introduced a new software architecture that fits the (memory, CPU, energy) constraints of low-end IoT devices, while being full-featured and easily extensible, thus more future-proof than state of the art. This work was published in ACM MobiSys'15 (IoT-Sys workshop), and released as open source code, integrated in the latest version of RIOT 2015-12. We have also designed a distributed test framework which supports advanced continuous integration techniques, allows for the integration of project contributors to volunteer hardware and software resources to the test system, and can function as a permanent distributed plugtest for network interoperability testing. This work was published in ACM MobiSys'15 (IoT-Sys workshop). Concerning IoT-lab, we have contributed to the completion of the design and the roll-out of IoT-lab testbeds in multiple sites in France and started deploying an additional one in Berlin. Description of completed work and design was published in IEEE IoT-WF'15.

6.2.2. Standards for Spontaneous Wireless Networks

Participant: Emmanuel Baccelli.

Within this activity, we have contributed to new network protocol standards for spontaneous wireless networking, applied to ad hoc networks and the Internet of Things. In particular, collaborating with Fraunhofer, we have published Directional Airtime Metric (DAT), a new wireless metric standard targeting wireless mesh networks. The standard is in the RFC editor's queue (which means the corresponding IETF standard, an RFC, will be published within weeks). Furthermore, collaborating with ARM and Sigma Designs, we published RFC 7733, which provides guidance in the configuration and use of protocols from the RPL protocol suite to implement the features required for control in building and home environments. In collaboration with various industrial partners, we have also published a number of other Internet drafts, including an analysis of the characteristics of multi-hop ad hoc wireless communication between interfaces in the context of IP networks, and an analysis of the challenges of information-centric networking in the Internet of Things.

6.3. Resource and Traffic Management

Traffic offloading; infrastructure deployment; opportunistic routing; traffic modeling; intermittently connected networks.

6.3.1. On the Interaction between Content Caching and Routing

Participants: Kolar Purushothama Naveen, Laurent Massoulié, Emmanuel Baccelli, Aline Carneiro Viana, Don Towsley.

Nowadays Internet users are mobile over 60% of their time online, and mobile data traffic is expected to increase by more than 60% annually to reach 15.9 exabytes per month by 2018. This evolution will likely incur durably congested wireless access at the edge despite progress in radio technologies. To alleviate congestion at the Internet edge, one promising approach is to target denser deployments of wireless access points. As a result, mobile users are potentially within radio reach of several access points (AP) from which content may be directly downloaded. In this context, distinct AP's can have very different bandwidth and memory capacities. Such differences raise the following question: When requests can be sent to several such access points, how to optimize performance through both load balancing and content replication?

In this work, we introduce formal optimization models to address this question, where bandwidth availability is represented via a cost function, and content availability is represented either by a cost function or a sharp constraint. For both formulations we propose dynamic caching and request assignment algorithms. Crucially our request assignment scheme is based on a server price signal jointly reflecting content and bandwidth availability. Using mean field approximation and Lyapunov functions techniques, we prove that our algorithms are optimal and stable in a limiting fluid regime with large arrival rates and content chunking. Through simulations we exhibit the efficacy of our request assignment strategy in comparison to the common practices

of assigning requests purely based on either bandwidth or content availability. Finally, using the popular LRU (Least Recently Used) strategy instead for cache replacements, we again demonstrate the superior performance of our request assignment strategies. This work was published in the ACM SIGCOMM'15 workshop on All Things Cellular.

6.3.2. *From Routine to Network Deployment for Data Offloading in Metropolitan Areas*

Participants: Eduardo Mucceli, Aline Carneiro Viana.

Smartphone sales are booming, nearly half billion were sold in 2011; more smartphones, more mobile data traffic, and Currently, 3G cellular networks in metropolitan areas are struggling to attend the recent boost up of mobile data consumption. Carefully deploying WiFi hotspots allow to maximize WiFi offloading and can both be cheaper than upgrade the cellular network structure and concede substantial improvement in the network capacity. In this context, in this work, we first propose a new way to map into a graph the *people behavior* (i.e., mobility context) in an urban scenario. Our proposed behavior-to-graph solution is simple, take into consideration the restrictions imposed by transportation modes to traffic demand, the space-time interaction between people and urban locations, and finally, is powerful to be used as input to any popular area identification problem (key points for an efficient network planning). Secondly, we propose a metric to identify locations more capable of providing coverage for people and consequently, more suitable for receiving hotspots. Deploying a small percentage of hotspots ranked by the herein proposed metric provides high percentages of coverage time for people moving around in the city. Using a real-life metropolitan trace, we show our routine-based strategy guarantees higher offload ratio than current approaches in the literature while using a realistic traffic model. This work, including new characterization results of the used trace and new analysis of space-traffic correlation, is under submission in a transaction.

6.3.3. *Mobile Data Traffic Modeling: Revealing Hidden Facets*

Participants: Eduardo Mucceli, Aline Carneiro Viana, Kolar Purushothama Naveen, Carlos Sarraute.

Smartphone devices provide today the best means of gathering users information about content consumption behavior on a large scale. In this context, the literature is rich in work studying and modeling users mobility, but little is publicly known about users content consumption patterns. The *understanding of users' mobile data traffic demands* is of fundamental importance when looking for solutions to manage the recent boost up of mobile data usage and to improve the quality of communication service provided. Hence, the definition of a *usage pattern* can allow telecommunication operators to better foresee future demanded traffic and consequently, to better (1) deploy data offloading hotspots or (2) timely plan network resources allocation and then, set subscription plans.

Using a large-scale dataset collected from a major 3G network in a big metropolitan area, in this work, we present the first detailed measurement-driven modeling of mobile data traffic usage of smartphone subscribers. Our main outcome is a synthetic measurement-based mobile data traffic generator, capable of imitating traffic-related activity patterns of different categories of subscribers and time periods of a routinary normal day in their lives. For this, we first characterize individual subscribers routinary behaviour, followed by the detailed investigation of subscribers' usage pattern (i.e., "when" and "how much" traffic is generated). Broadly, our observations bring important insights into network resource usage. We then classify the subscribers into six distinct profiles according to their usage pattern and model these profiles according to two different journey periods: peak and non-peak hours. We show that the synthetic trace generated by our data traffic model consistently imitates different subscriber profiles in two journey periods, when compared to the original dataset. We discuss relevant issues in traffic demands and describe implications in network planning and privacy. This work, including a new characterization results of the used trace, including analysis correlating age and gender to traffic demands, as well as new profiling results, is under submission in a transaction.

6.3.4. *Data Delivery in Opportunistic and Intermittently Connected Networks*

Participants: Ana Cristina Vendramin, Anelise Munaretto, Myriam Delgado, Aline Carneiro Viana, Mauro Fonseca.

The pervasiveness of computing devices and the emergence of new applications and cloud services are factors emphasizing the increasing need for adaptive networking solutions. In most cases, this adaptation requires the design of interdisciplinary approaches as those inspired by nature, social structures, games, and control systems. The approach presented in this work brings together solutions from different, yet complementary domains, i.e., networking, artificial intelligence, and complex networks, and is aimed at addressing the problem of efficient data delivery in intermittently connected networks.

As mobile devices become increasingly powerful in terms of communication capabilities, the appearance of opportunistic and intermittently connected networks referred to as Delay Tolerant Networks (DTNs) is becoming a reality. In such networks, contacts occur opportunistically in corporate environments such as conferences sites, urban areas, or university campuses. Understanding node mobility is of fundamental importance in DTNs when designing new communication protocols that consider opportunistic encounters among nodes. This work proposes the Cultural Greedy Ant (CGrAnt) protocol to solve the problem of data delivery in opportunistic and intermittently connected networks. CGrAnt is a hybrid Swarm Intelligence-based forwarding protocol designed to address the dynamic and complex environment of DTNs. CGrAnt is based on: (1) Cultural Algorithms (CA) and Ant Colony Optimization (ACO) and (2) operational metrics that characterize the opportunistic social connectivity between wireless users. The most promising message forwarders are selected via a greedy transition rule based on local and global information captured from the DTN environment. Using simulations, we first analyze the influence of the ACO operators and CA knowledge on the CGrAnt performance. We then compare the performance of CGrAnt with the PROPHET and Epidemic protocols (two well known related protocols in the literature) under varying networking parameters. The results show that CGrAnt achieves the highest delivery ratio (gains of 99.12% compared with PROPHET and 40.21% compared with Epidemic) and the lowest message replication (63.60% lower than PROPHET and 60.84% lower than Epidemic). This work is under submission to an international journal.

6.3.5. Designing Adaptive Replication Schemes in Distributed Content Delivery Networks

Participants: Mathieu Leconte, Marc Lelarge, Laurent Massoulié.

In a paper published at the ITC'15 conference, we address the problem of content replication in large distributed content delivery networks, composed of a data center assisted by many small servers with limited capabilities and located at the edge of the network. We aim at optimizing the placement of contents on the servers to offload the data center as much as possible. We model the sub-system constituted by the small servers as a loss network, each loss corresponding to a request to the data center. Based on large system / storage behavior, we obtain an asymptotic formula for the optimal replication of contents and propose adaptive schemes to attain it by reacting to losses, as well as faster algorithms which can react before losses occur. We show through simulations that our adaptive schemes outperform significantly standard replication strategies both in terms of loss rates and adaptation speed.

6.3.6. Vehicular Network under a Social Perception

Participants: Felipe D. Cunha, Aline Carneiro Viana, Raquel A. F. Mini, Antonio A.f. Loureiro.

Vehicular Mobility is strongly influenced by the speed limits, destinations, traffic conditions, period of the day, and direction of the public roads. At the same time, the driver's behavior produces great influences in vehicular mobility. People tend to go to the same places, at the same day period, through the same trajectories, which lead them to the appearance of driver's daily routines. These routines lead us to the study of mobility in VANETs under a social perspective and to investigate how effective is to explore social interactions in this kind of network. In this work, we thus characterize and evaluate social properties of a realistic vehicular trace found in literature. Our aim is to study the vehicles' mobility in accordance to social behaviors. Social metrics are computed and the obtained results are compared to random graphs. With our analysis, we could verify the existence of regularity and common interests among the drivers in vehicular networks.

After having identified routine in vehicles mobility patterns and their correlation with the period of the day, we then leverage the identified social aspects to design a *Socially Inspired Broadcast Data Dissemination* for VANETs. We claim that protocols and applications designed for Vehicular Ad Hoc Networks need to adapt to

vehicles routines in order to provide better services. With this issue in mind, we designed a data dissemination solution for these networks that considers the daily road traffic variation of large cities and the relationship among vehicles. The focus of our approach is to select the best vehicles to rebroadcast data messages according to social metrics, in particular, the clustering coefficient and the node degree. Moreover, our solution is designed in such a way that it is completely independent of the perceived road traffic density. Simulation results show that, when compared to related protocols, our proposal provides better delivery guarantees, reduces the network overhead and possesses an acceptable delay.

6.3.7. Design and Analysis of an Efficient Friend-to-Friend Content Dissemination System

Participants: Kanchana Thilakarathna, Aline Carneiro Viana, Aruna Seneviratne, Henrik Petander.

In this work, we focus on dissemination of content for delay tolerant applications/services, (i.e. content sharing, advertisement propagation, etc.) where users are geographically clustered into communities. Due to emerging security and privacy concerns, majority of users are becoming more reluctant to interact with strangers and are only willing to share information/content with the users who are previously identified as friends. As a result, despite its promise, opportunistic communications systems have not been widely adopted. In addition, in this environment, opportunistic communication will not be effective due to the lack of known friends within the communication range. We thus propose a novel architecture which combines the advantages of distributed decentralized storage and opportunistic communications. The proposed system addresses the trust and privacy concerns of opportunistic communications systems, and enables the provision of efficient distributed mobile social networking services. We exploit the fact that users will trust their friends, and the friends will help in disseminating content by temporarily storing and forwarding content. This can be done by replicating content on friends' devices who are likely to consume that content and provide the content to other friends when the device has access to low cost networks. The fundamental challenge then is to minimize the number of replicas, to ensure high and timely availability. We provide a formal definition of this content replication problem, and show that it is NP hard. Then, we propose a community based greedy heuristic algorithm with novel dynamic centrality metrics that replicates the content on a minimum number of friends' devices, and maximizes the availability of content. Using both real world and synthetic traces, we validate effectiveness of the proposed scheme. In addition, we demonstrate the practicality of the the proposed system, through an implementation on Android smartphones. This work is under submission in an international transaction.

6.3.8. Telling Apart Social and Random Relationships in Dynamic Networks

Participants: Pedro Olmo Vaz de Melo, Aline Carneiro Viana, Marco Fiore, Katia Jaffrès-Runser, Frédéric Le Mouël, Antonio A. F. Loureiro, Lavanya Addepalli, Guangshuo Chen.

Recent studies have analyzed data generated from mobile individuals in urban regions, such as cab drivers or students in large campuses. Particular attention has been paid to the dynamics of user movement, whose real-world complexity cannot be fully captured through synthetic models. Indeed, understanding user mobility is of fundamental importance when designing new communication protocols that exploit opportunistic encounters among users. In this case, the problem mainly lies in correctly forecasting future contacts. To that end, the regularity of daily activities comes in handy, as it enforces periodic (and thus predictable) space-time patterns in human mobility. Although human behavior is characterized by an elevated rate of regularity, random events are always possible in the routines of individuals. Those are hardly predictable situations that deviate from the regular pattern and are unlikely to repeat in the future.

We argue that the ability to accurately spot random and social relationships in dynamic networks is essential to network applications that rely on a precise description of human routines, such as recommendation systems, forwarding strategies and opportunistic dissemination protocols. We thus propose a strategy to analyze users' interactions in mobile networks where users act according to their interests and activity dynamics. Our strategy, named *Random rElationship ClAssifier sTrategy (RECAST)*, allows classifying users' wireless interactions, separating random interactions from different kinds of social ties. To that end, RECAST observes how the real system differs from an equivalent one where entities' decisions are completely random. We evaluate the effectiveness of the RECAST classification on five real-world user contact datasets collected in diverse networking contexts. Our analysis unveils significant differences among the dynamics of users' wireless

interactions in the datasets, which we leverage to unveil the impact of social ties on opportunistic routing. We show that, for such specific purpose, the relationships inferred by classifier are more relevant than, e.g., self-declared friendships on Facebook. This work was published in 2015 at the Performance Evaluation Elsevier Journal [9].

KERDATA Project-Team

7. New Results

7.1. Efficient data management for hybrid and multi-site clouds

7.1.1. *JetStream: enabling high-throughput live event streaming on multi-site clouds*

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

Scientific and commercial applications operate nowadays on tens of cloud datacenters around the globe, following similar patterns: they aggregate monitoring or sensor data, assess the QoS or run global data mining queries based on inter-site event stream processing. Enabling fast data transfers across geographically distributed sites allows such applications to manage the continuous streams of events in real time and quickly react to changes. However, traditional event processing engines often consider data resources as second-class citizens and support access to data only as a side-effect of computation (i.e. they are not concerned by the transfer of events from their source to the processing site). This is an efficient approach as long as the processing is executed in a single cluster where nodes are interconnected by low latency networks. In a distributed environment, consisting of multiple datacenters, with orders of magnitude differences in capabilities and connected by a WAN, this will undoubtedly lead to significant latency and performance variations.

This is namely the challenge we addressed this year by proposing JetStream [15], a high performance batch-based streaming middleware for efficient transfers of events between cloud datacenters. JetStream is able to self-adapt to the streaming conditions by modeling and monitoring a set of context parameters. It further aggregates the available bandwidth by enabling multi-route streaming across cloud sites, while at the same time optimizing resource utilization and increasing cost efficiency. The prototype was validated on tens of nodes from US and Europe datacenters of the Windows Azure cloud with synthetic benchmarks and a real-life application monitoring the ALICE experiment at CERN. The results show a $3\times$ increase of the transfer rate using the adaptive multi-route streaming, compared to state of the art solutions.

7.1.2. *Multi-site metadata management for geographically distributed cloud workflows*

Participants: Luis Eduardo Pineda Morales, Alexandru Costan, Gabriel Antoniu.

With their globally distributed datacenters, clouds now provide an opportunity to run complex large-scale applications on dynamically provisioned, networked and federated infrastructures. However, there is a lack of tools supporting data-intensive applications (e.g. scientific workflows) on virtualized IaaS or PaaS systems across geographically distributed sites. As a relevant example, data-intensive scientific workflows struggle in leveraging such distributed cloud platforms. For instance, scientific workflows which handle many small files can easily saturate state-of-the-art distributed filesystems based on centralized metadata servers (e.g., HDFS, PVFS).

In [22], we explore several alternative design strategies to efficiently support the execution of existing workflow engines across multi-site clouds, by reducing the cost of metadata operations. These strategies leverage workflow semantics in a 2-level metadata partitioning hierarchy that combines distribution and replication. The system was validated on the Microsoft Azure cloud across 4 EU and US datacenters. The experiments were conducted on 128 nodes using synthetic benchmarks and real-life applications. We observe as much as 28% gain in execution time for a parallel, geo-distributed real-world application (Montage) and up to 50% for a metadata-intensive synthetic benchmark, compared to a baseline centralized configuration.

7.1.3. *Understanding the performance of Big Data platforms in hybrid and multi-site clouds*

Participants: Roxana-Ioana Roman, Ovidiu-Cristian Marcu, Alexandru Costan, Gabriel Antoniu.

Recently, hybrid multi-site big data analytics (that combines on-premise with off-premise resources) has gained increasing popularity as a tool to process large amounts of data on-demand, without additional capital investment to increase the size of a single datacenter. However, making the most out of hybrid setups for big data analytics is challenging because on-premise resources can communicate with off-premise resources at significantly lower throughput and higher latency. Understanding the impact of this aspect is not trivial, especially in the context of modern big data analytics frameworks that introduce complex communication patterns and are optimized to overlap communication with computation in order to hide data transfer latencies. This year we started to work on a study that aims to identify and explain this impact in relationship to the known behavior on a single cloud.

A first step towards this goal consisted of analysing a representative big data workload on a hybrid Spark setup [24]. Unlike previous experience that emphasized low end-impact of network communications in Spark, we found significant overhead in the shuffle phase when the bandwidth between the on-premise and off-premise resources is sufficiently small. We plan to continue this study by investigating additional parameters at a finer grain and adding new platforms, like Apache Flink.

7.2. Optimizing Map-Reduce

7.2.1. Chronos: failure-aware scheduling in shared Hadoop clusters

Participants: Orçun Yildiz, Shadi Ibrahim, Gabriel Antoniu.

Hadoop emerged as the de facto state-of-the-art system for MapReduce-based data analytics. The reliability of Hadoop systems depends in part on how well they handle failures. Currently, Hadoop handles machine failures by re-executing all the tasks of the failed machines (i.e., executing recovery tasks). Unfortunately, this elegant solution is entirely entrusted to the core of Hadoop and hidden from Hadoop schedulers. The unawareness of failures therefore may prevent Hadoop schedulers from operating correctly towards meeting their objectives (e.g., fairness, job priority) and can significantly impact the performance of MapReduce applications.

In [23], we propose Chronos, a failure-aware scheduling strategy that enables an early yet smart action for fast failure recovery while operating within a specific scheduler objective. Chronos takes an early action rather than waiting an uncertain amount of time to get a free slot (thanks to our preemption technique). Chronos embraces a smart selection algorithm that returns a list of tasks that need to be preempted in order to free the necessary slots to launch recovery tasks immediately. This selection considers three criteria: the progress scores of running tasks, the scheduling objectives, and the recovery tasks input data locations. In order to make room for recovery tasks rather than waiting an uncertain amount of time, a natural solution is to kill running tasks in order to create free slots. Although killing tasks can free the slots easily, it wastes the work performed by the killed tasks. Therefore, we present the design and implementation of a novel work-conserving preemption technique that allows pausing and resuming both map and reduce tasks without resource wasting and with little overhead.

We demonstrate the utility of Chronos by combining it with two state-of-the-art Hadoop schedulers: Fifo and Fair schedulers. The experimental results show that Chronos achieves almost optimal data locality for the recovery tasks and reduces the job completion times by up to 55% over state-of-the-art schedulers. Moreover, Chronos recovers to a correct scheduling behavior after failure detection within only a couple of seconds.

7.2.2. On the usability of shortest remaining time first policy in shared Hadoop clusters

Participants: Nathanaël Cherièr, Shadi Ibrahim.

A practical problem facing the Hadoop community is how to reduce job makespans by reducing job waiting times and execution times. Previous Hadoop schedulers have focused on improving job execution times, by improving data locality but not considering job waiting times. Even worse, enforcing data locality according to the job input sizes can be inefficient: it can lead to long waiting times for small yet short jobs when sharing the cluster with jobs with smaller input sizes but higher execution complexity.

We have introduced hSRTF [16], an adaption of the well-known Shortest Remaining Time First scheduler (i.e., SRTF) in shared Hadoop clusters. hSRTF embraces a simple model to estimate the remaining time of a job and a preemption primitive (i.e., kill) to free the resources when needed. We have implemented hSRTF and performed extensive evaluations with Hadoop on the Grid'5000 testbed. The results show that hSRTF can significantly reduce the waiting times of small jobs and therefore improves their make-spans, but at the cost of a relatively small increase in the make-spans of large jobs. For instance, a time-based proportional share mode of hSRTF (i.e., hSRTF-Pr) speeds up small jobs by (on average) 45% and 26% while introducing a performance degradation for large jobs by (on average) 10% and 0.2% compared to Fifo and Fair schedulers, respectively.

7.2.3. A Performance evaluation of Hadoop's schedulers under failures

Participants: Shadi Ibrahim, Gabriel Antoniu.

Recently, Hadoop has not only been used for running single batch jobs but it has also been optimized to simultaneously support the execution of multiple jobs belonging to multiple concurrent users. Several schedulers (i.e., Fifo, Fair, and Capacity schedulers) have been proposed to optimize locality executions of tasks but do not consider failures, although, evidence in the literature shows that faults do occur and can probably result in performance problems.

In [19], we have designed a set of experiments to evaluate the performance of Hadoop under failure when applying several schedulers (i.e., explore the conflict between job scheduling, exposing locality executions, and failures). Our results reveal several drawbacks of current Hadoop's mechanism in prioritizing failed tasks. By trying to launch failed tasks as soon as possible regardless of locality, it significantly increases the execution time of jobs with failed tasks, due to two reasons: 1) available resources might not be freed up as quickly as expected and 2) failed tasks might be re-executed on machines with no data on it, introducing extra cost for data transferring through network, which is normally the most scarce resource in today's datacenters.

Our preliminary study with Hadoop not only helps us to understand the interplay between fault-tolerance and job scheduling, but also offers useful insights into optimizing the current schedulers to be more efficient in case of failures.

7.2.4. Kvasir: empowering Hadoop with knowledge

Participants: Nathanaël Cherièr, Shadi Ibrahim.

Most of Hadoop schedulers are based on homogeneity hypotheses about the jobs and the nodes and therefore strongly rely on the location of the input data when scheduling tasks. However, our study revealed that Hadoop is a highly dynamic environment (e.g., variation in task duration within a job and across different jobs). Even worse, clouds are multi-tenant environments which in turn introduce more heterogeneity and dynamicity in Hadoop clusters. As a result, relying on static knowledge (i.e. data location) may lead to wrong scheduling decisions.

We have developed a new scheduling framework for Hadoop, named Kvasir. Kvasir aims to provide an up-to-date knowledge that reflects the dynamicity of the environment while being light-weight and performance-oriented. The utility of Kvasir is demonstrated by the implementation of several schedulers including Fifo, Fair, and SRTF schedulers.

7.3. Energy-aware data management in clouds and HPC

7.3.1. On understanding the energy impact of speculative execution in Hadoop

Participants: Tien Dat Phan, Shadi Ibrahim, Gabriel Antoniu, Luc Bougé.

Hadoop emerged as an important system for large-scale data analysis. Speculative execution is a key feature in Hadoop that is extensively leveraged in clouds: it is used to mask slow tasks (i.e., stragglers) — resulted from resource contention and heterogeneity in clouds — by launching speculative task copies on other machines. However, speculative execution is not cost-free and may result in performance degradation and extra resource and energy consumption. While prior literature has been dedicated to improving stragglers detection to cope with the inevitable heterogeneity in clouds, little work is focusing on understanding the implications of speculative execution on the performance and energy consumption in Hadoop cluster.

In [21], we have designed a set of experiments to evaluate the impact of speculative execution on the performance and energy consumption of Hadoop in homogeneous and heterogeneous environments. Our studies reveal that speculative execution may sometimes reduce, sometimes increase the energy consumption of Hadoop clusters. This strongly depends on the reduction in the execution time of MapReduce applications and on the extra power consumption introduced by speculative execution. Moreover, we show that the extra power consumption varies among applications and is contributed to by three main factors: the duration of speculative tasks, the idle time, and the allocation of speculative tasks. To the best of our knowledge, our work provides the first deep look into the energy efficiency of speculative execution in Hadoop.

7.3.2. *On the energy footprint of I/O management in Exascale HPC systems*

Participants: Orçun Yildiz, Matthieu Dorier, Shadi Ibrahim, Gabriel Antoniu.

The advent of unprecedentedly scalable yet energy hungry Exascale supercomputers poses a major challenge in sustaining a high performance-per-watt ratio. With I/O management acquiring a crucial role in supporting scientific simulations, various I/O management approaches have been proposed to achieve high performance and scalability. However, the details of how these approaches affect energy consumption have not been studied yet.

Therefore, we have explored how much energy a supercomputer consumes while running scientific simulations when adopting various I/O management approaches. In particular, we closely examined three radically different I/O schemes including time partitioning, dedicated cores, and dedicated nodes. To do so, we implemented the three approaches within the Damaris I/O middleware and performed extensive experiments with one of the target HPC applications of the Blue Waters sustained-petaflop supercomputer project: the CM1 atmospheric model.

Our experimental results obtained on the French Grid'5000 platform highlighted the differences among these three approaches and illustrate in which way various configurations of the application and of the system can impact performance and energy consumption. Considering that choosing the most energy-efficient approach for a particular simulation on a particular machine can be a daunting task, we provided a model to estimate the energy consumption of a simulation under different I/O approaches. Our proposed model gives hints to pre-select the most energy-efficient I/O approach for a particular simulation on a particular HPC system and therefore provides a step towards energy-efficient HPC simulations in Exascale systems.

We validated the accuracy of our proposed model using a real-life HPC application (CM1) and two different clusters provisioned on the Grid'5000 testbed. The estimated energy consumptions are within 5.7% of the measured ones for all I/O approaches.

7.3.3. *Exploring energy-consistency trade-offs in cloud storage systems and beyond*

Participants: Mohammed-Yacine Taleb, Shadi Ibrahim, Gabriel Antoniu, Luc Bougé.

Apache Cassandra is an open-source cloud storage system that offers multiple types of operation-level consistency including eventual consistency with multiple levels of guarantees and strong consistency. It is being used by many datacenter applications (e.g., Facebook and AppScale). Most existing research efforts have been dedicated to exploring trade-offs such as: consistency vs. performance, consistency vs. latency and consistency vs. monetary cost. In contrast, a little work is focusing on the consistency vs. energy trade-off. As power bills have become a substantial part of the monetary cost for operating a datacenter, we aim to provide a clearer understanding of the interplay between consistency and energy consumption.

In [17], a series of experiments have been conducted to explore the implication of different factors on the energy consumption in Cassandra. Our experiments have revealed a noticeable variation in the energy consumption depending on the consistency level. Furthermore, for a given consistency level, the energy consumption of Cassandra varies with the access pattern and the load exhibited by the application. This further analysis indicated that the uneven distribution of the load amongst different nodes also impacts the energy consumption in Cassandra. Finally, we experimentally compared the impact of four storage configuration and data partitioning policies on the energy consumption in Cassandra: interestingly, we achieve 23% energy saving when assigning 50% of the nodes to the hot pool for the applications with moderate ratio of reads and writes, while applying eventual (quorum) consistency.

This study points to opportunities for future research on consistency-energy trade-offs and offers useful insight into designing energy-efficient techniques for cloud storage systems. This work was done in collaboration with Houssein-Eddine Chihoub (LIG lab, Grenoble) and María Pérez (UPM, Madrid).

Recently, we have been looking at in-memory storage systems. In particular, we are investigating the current replication schemes, data placement strategies and consistency models which are used in in-memory storage systems. Next, an empirical study will be performed to analyze the potential impact of the aforementioned issues on energy consumption. At this point, we are working with RAMCloud.

7.3.4. Governing energy consumption in Hadoop through CPU frequency scaling: an analysis

Participants: Tien Dat Phan, Shadi Ibrahim, Gabriel Antoniu.

In [12], we studied the impact of different existing DVFS (*Dynamic Voltage and Frequency Scaling*) governors (i.e., performance, powersave, on-demand, conservative and userspace) on Hadoop's performance and power efficiency. Interestingly, our experimental results reported not only a noticeable variation of the power consumption and performance with different applications and under different governors, but also demonstrate the opportunity to achieve a better tradeoff between performance and power consumption.

The primary contributions of this work are as follows: (1) it provides an overview of the state-of-the-art techniques for energy-efficiency in Hadoop; (2) it discusses and demonstrates the need for exploiting DVFS techniques for energy reduction in Hadoop; (3) it experimentally demonstrates that MapReduce applications experience variations in performance and power consumption under different CPU frequencies and also under different governors. A micro-analysis section is provided to explain this variation and its cause; (4) it illustrates in practice how the behavior of different governors influences the execution of MapReduce applications and how it shapes the performance of the entire cluster; (5) it also brings out the differences between these governors and CPU frequencies and shows that they are not only sub-optimal for different applications but also sub-optimal for different stages of MapReduce execution; (6) it demonstrates that achieving better energy efficiency in Hadoop cannot be done simply by tuning the governor parameters, nor through a naive coarse-grained tuning of the CPU frequencies or the governors according to the running phase (i.e., map phase or reduce phase).

7.4. Scalable I/Os: visualization and processing

7.4.1. Modeling and predicting I/O patterns of large-scale simulations

Participants: Matthieu Dorier, Shadi Ibrahim, Gabriel Antoniu.

The increasing gap between the computation performance of post-petascale machines and the performance of their I/O subsystem has motivated many I/O optimizations including prefetching, caching, and scheduling. In order to further improve these techniques, modeling and predicting spatial and temporal I/O patterns of HPC applications as they run has become crucial. Our work in this context focuses on Omnisc'IO, an approach that builds a grammar-based model of the I/O behavior of HPC applications and uses it to predict when future I/O operations will occur, and where and how much data will be accessed. To infer grammars, Omnisc'IO is based on StarSequitur, a novel algorithm extending Nevill-Manning's Sequitur algorithm [11]. Omnisc'IO is transparently integrated into the POSIX and MPI I/O stacks and does not require any modification in applications or higher-level I/O libraries. It works without any prior knowledge of the application and converges to accurate predictions of any N future I/O operations within a couple of iterations. Its implementation is efficient in both computation time and memory footprint.

7.4.2. In situ analysis and visualization workflows

Participants: Matthieu Dorier, Lokman Rahmani, Gabriel Antoniu.

In situ visualization has been proposed in the past few years to couple running simulations with parallel visualization and analysis tools. While many parallel visualization tools now provide in situ visualization capabilities, the trend has been to feed such tools with what previously was large amounts of unprocessed output data and let them render everything at the highest possible resolution. This leads to an increased run time of simulations that still have to complete within a fixed-length job allocation. In this work, we tackle the challenge of enabling in situ visualization under performance constraints. Our approach shuffles data across processes according to its content and filters out part of it in order to feed a visualization pipeline with only a reorganized subset of the data produced by the simulation. Our framework monitors its own performance and reconfigures itself dynamically to achieve the best possible visual fidelity within predefined performance constraints. Experiments on the Blue Waters supercomputer with the CM1 simulation show that our approach enables a $5\times$ speedup and is able to meet performance constraints.

7.5. Scalable storage for data-intensive applications

7.5.1. *OverFlow: multi-site aware Big Data management for scientific workflows on clouds*

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

The global deployment of cloud datacenters is enabling large-scale scientific workflows to improve performance and deliver fast responses. This unprecedented geographical distribution of the computation is doubled by an increase in the scale of the data handled by such applications, bringing new challenges related to the efficient data management across sites. High throughput, low latencies or cost-related trade-offs are just a few concerns for both cloud providers and users when it comes to handling data across datacenters. Existing solutions are limited to cloud-provided storage, which offers low performance based on rigid cost schemes. In turn, workflow engines need to improvise substitutes, achieving performance at the cost of complex system configurations, maintenance overheads, reduced reliability and reusability.

In [14], we introduced OverFlow, a uniform data-management system for scientific workflows running across geographically distributed sites, aiming to reap economic benefits from this geo-diversity. Our solution is environment-aware, as it monitors and models the global cloud infrastructure, offering high and predictable data-handling performance for transfer cost and time, within and across sites. OverFlow proposes a set of pluggable services, grouped in a data-scientist cloud kit. They provide the applications with the possibility to monitor the underlying infrastructure, to exploit smart data compression, deduplication and geo-replication, to evaluate data-management costs, to set a tradeoff between money and time, and optimize the transfer strategy accordingly. The system was validated on the Microsoft Azure cloud across its 6 EU and US datacenters. The experiments were conducted on hundreds of nodes using synthetic benchmarks and real-life bio-informatics applications (A-Brain, BLAST). The results show that our system is able to model the cloud performance accurately and to leverage this for efficient data dissemination, being able to reduce the monetary costs and transfer time by up to 3 times.

7.5.2. *Efficient transactional storage for data-intensive applications*

Participants: Pierre Matri, Alexandru Costan, Gabriel Antoniu.

As the computational power used by large-scale applications increases, the amount of data they need to manipulate tends to increase as well. A wide range of such applications require robust and flexible storage support for atomic, durable and concurrent transactions. Historically, databases have provided the *de facto* solution to transactional data management, but they have forced applications to drop control over data layout and access mechanisms, while remaining unable to meet the scale requirements of Big Data. More recently, key-value stores have been introduced to address these issues. However, this solution does not provide transactions, or only restricted transaction support, constraining users to carefully coordinate access to data in order to avoid race conditions, partial writes, overwrites, and other hard problems that cause erratic behavior.

We argue that there is a gap between existing storage solutions and application requirements that limits the design of transaction-oriented data-intensive applications. We have started working on a prototype of a massively parallel distributed transactional blob storage system, aiming to fill this gap.

MADYNES Project-Team

7. New Results

7.1. Monitoring

7.1.1. Anonymous networks monitoring

Participants: Thibault Cholez [contact], Isabelle Chrisment, Olivier Festor.

In 2015, we pursued our collaboration with Juan Pablo Timpanaro a former team's PhD student and published a new paper [47] on the I2P anonymous network (<http://i2p2.de>). More precisely, we monitored I2P's decentralised directory, known as the netDB, and produced two contributions. On the one hand, we conducted arguably the first *churn* study of the I2P network, showing that I2P users are more stable than non-anonymous peer-to-peer users. On the other hand, we analysed the design of the netDB and compared it against the popular KAD design, demonstrating that the former is more vulnerable to different attacks, specially to Eclipse attacks, which can be mitigated by applying some safer design choices of the latter. We lately showed the positive impact on performance of including KAD's DHT configuration into the netDB in terms of bandwidth, storage and messages overhead.

7.1.2. Smartphone usage monitoring

Participants: Vassili Rivron [contact], Mohammad Irfan Khan, Simon Charneau [Inria], Isabelle Chrisment.

In [39] we presented some results from our study based on a combination of crowdsending and survey. We discussed some technical problems we faced and some lessons learned during our crowdsensing experiment. Furthermore we showed how information regarding social context can be used for better interpretation of crowdsensed data. Next we selected some questions from the multiple choice survey questionnaire and combined the responses with crowdsensed data to analyze users' perception about their smartphone usage and discussed cognitive factors associated with reporting information on questionnaires. Moreover we showed that combining sensing with survey can improve both the techniques and the combination has important use cases such as helping users to have a better understanding and control of their technology usage.

7.1.3. Active Monitoring

Participants: Abdelkader Lahmadi [contact], Jérôme François, Valentin Giannini, Frederic Beck [LHS], Bertrand Wallrich [LHS].

The main motivation of this work was to assess the exposition of industrial systems in the Internet, especially by measuring how many SCADA systems are accessible. To do so, we built an IPv4 methodology which is able to scan the entire IPv4 address space by maximizing the distance between consecutive IP addresses. It thus avoids colateral effect of overloading targeted networks and being blacklisted. We thus extend the Zmap tool (zmap.io) by also including other functionalities such as distributed scans, indexation and visualisation of the results [63]. First experiences have been performed and are under evaluation.

7.1.4. Sensor networks monitoring

Participants: Rémi Badonnel, Isabelle Chrisment, Olivier Festor, Abdelkader Lahmadi [contact], Anthea Mayzaud.

This year, our work on security-oriented monitoring has been centered on building a distributed architecture that supports passive monitoring in the Internet of Things using the RPL protocol [37]. A particular interest has been given to advanced metering infrastructure (AMI) networks, where higher order devices are expected to form the backbone infrastructure, to which more constrained nodes would connect. Our distributed architecture exploits the capabilities of these higher order devices to perform network monitoring tasks, and takes benefits from properties inherent to that protocol, such as DODAG building and multi-instance routing mechanisms, in order to passively monitor the environment with a minimal impact on constrained nodes.

We have also consolidated our taxonomy on security attacks in these networks [8]. In addition, we have pursued our work on topological inconsistency attacks [9]. It is evident from the experiments that we have conducted that mitigating such attacks is critical to avoid channel congestion and high resource usage. Our initial adaptive threshold (AT) strategy to mitigate the effects of such attacks has been further improved. The new strategy dynamically takes into account network characteristics in order to infer an appropriate threshold for counteracting these attacks.

7.2. Security

7.2.1. Security analytics

Participants: Jérôme François [contact], Abdelkader Lahmadi, Manobala Nirmala, Vincent Noyalet.

During the year 2015, we have extended our monitoring platform dedicated to Android environments [69] with more analytics features. The monitoring platform is dedicated to the collection, storage, analysis and visualization of logs and network flow data of mobile applications. The platform relies on a set of on-device probes to monitor network and system activities of these applications. The data are collected from these probes and parsed through generic and flexible collectors relying on Flume agents that we have adapted and extended. We are storing the collected data using a column oriented Hbase storage engine (Hadoop database). Finally, after being parsed, the data are made available within the Elasticsearch engine to search and visualize them using the Kibana tool. We have also presented the building blocks of the platform in a lab session within the conference AIMS 2015 [70].

We have also maintained an IETF draft [75] to promote a standardization effort towards the extension of IP Flow-based monitoring with geographic information. Associating Flow information with their measurement geographic locations will enable security applications to detect anomalous activities. In the case of mobile devices, the characterization of communication patterns using only time and volume is not enough to detect unusual location-related communication patterns.

Besides, we looked at aggregating flows collected at the High Security Lab since a single attack is represented by multiple flows. For example, a DDoS or a scan is a sequence of similar parallel flows coming from the same or distributed machines. As attacks occur very frequently and even at the same time, grouping flows occurring in a pre-defined time window is not a valid approach. Two approaches have been investigated and are actually dependent of the sources of collected flows. First, we analyzed collected Netflow data from the Darknet which is basically a sinkhole without any services running or announced. Hence, all traffic is considered as abnormal and is limited to a set of predefined attacks. Indeed, since no packets can be sent back, complex attacks with different steps cannot be caught. Therefore, scanning, flooding-based denial-of-service and backscatter are the main types of anomalies we can observe. Flows are thus grouped and labeled regarding certain criteria (common IP addresses/subnets, ports, co-occurrence) thanks to a pre-established decision process [58]. The final goal was to compare data collected in Nancy and in Tokyo. Secondly, we assume flow data without specific knowledge about the type of traffic it embeds. In such a case, the goal is to automatically extract recurrent patterns. The initial approach consisted in representing flows as nodes in a graph and linking them when sharing some properties (IP addresses, ports). Major subsequent problems have been faced like indexation, split flows in multiple files and visualization [59].

7.2.2. Management of HTTPS traffic

Participants: Thibault Cholez [contact], Shbair Wazen, Jérôme François, Isabelle Chrisment.

We previously investigated the latest technique for HTTPS traffic filtering that is based on the Server Name Indication (SNI) field of TLS and which has been recently implemented in many firewall solutions. We showed that SNI has two weaknesses, regarding (1) backward compatibility and (2) multiple services using a single certificate and we implemented a proof of concept of these vulnerabilities as a web browser extension (Escape). This work was published in the IFIP/IEEE IM'15 conference [44].

This led us to the development of new reliable methods to investigate the increasing number of HTTPS traffic that may hold security breaches but without relying on decryption at any step, in order to respect users' privacy (no HTTPS proxy). Many approaches already identify the main type of an application (Web, P2P, SSH,...) running in secure tunnels, and others identify a couple of specific encrypted web pages through website fingerprinting.

In this context, we developed a better technique to precisely identify the services run within HTTPS connections, i.e. to name the services, without relying on specific header fields that can be easily altered. We have defined dedicated features for HTTPS traffic that are used as input for a multi-level identification framework based on machine learning algorithms. Our evaluation based on real traffic shows that we can identify encrypted web services with a high accuracy. This work will be published next year in the IFIP/IEEE Network Operations and Management Symposium (NOMS 2016).

7.2.3. Configuration security automation

Participants: Rémi Badonnel [contact], Gaetan Hurel, Abdelkader Lahmadi, Olivier Festor.

Our work during year 2015 was mainly focused on the orchestration of security functions in the context of mobile smart environments [35]. Most of current security approaches for these environments are provided in the form of applications or packages to be directly installed on the devices themselves. Such approaches may be qualified as on-device. However, on-device approaches generally induce significant local resource consumption leading to the significant reduction of battery lifetime. In the meantime, current cloud-based approaches for mobile security attempt to deal with this issue by offloading most of the workload on a remote server, but may introduce significant additional latency. In that context, we have pursued the efforts on our strategy for dynamically outsourcing and composing security functions in the cloud, considering software-defined networking. The architecture relies on a set of security functions that are activated, configured and orchestrated according to the current contexts and risks, while a dedicated modelling has been introduced for supporting the evaluation of security compositions and their properties. The chaining of security functions is performed dynamically in order to fit with the security requirements of mobile devices at runtime. In particular, we have proposed in [35] to analyze and cluster applications running on the mobile devices based on their network behaviors, in order to drive the selection and deployment of adequate security compositions that may be fully outsourced or split between in-cloud and on-device.

We have also investigated in [23] to what extent security automation, more specifically in the context of vulnerability management, might be supported by conceptual knowledge discovery. The intended extension might be a mean to cope with the increasing dynamics and complexity of networked environments. Most current security solutions still seem to work under certain boundaries that prevent them to act intelligently and flexibly, i.e. strictly stucked to the available security information in order to analyze, report and eventually remediate found problems. Our purpose is to exploit methods and techniques coming from formal concept and knowledge discovery in databases, in order to provide high-level automation based on mechanisms capable of understanding, reasoning about, and anticipating the surrounding environment and its vulnerabilities.

7.3. Experimentation, Emulation, Reproducible Research

This section covers our work on experimentation on testbeds (mainly Grid'5000), on emulation (mainly on Distem), and on Reproducible Research.

7.3.1. Grid'5000 design and evolutions

Participants: Jérémie Gaidamour, Arthur Garnier, Lucas Nussbaum [contact], Clément Parisot.

The team was again heavily involved in the evolutions and the governance of the Grid'5000 testbed.

In the context of ADT LAPLACE, Jérémie Gaidamour adapted and configured the CiGri middleware on Grid'5000. CiGri enables the execution of large campaigns of *best-effort* jobs (low priority, interruptible jobs). It is expected that this work will allow the remaining free time slots to be filled by tasks from other research communities such as natural language processing.

Jérémie Gaidamour also greatly improved *stats5k*, our tool to generate metrics about the testbed (usage, resources availability, etc.), available at <https://intranet.grid5000.fr/stats/>.

Arthur Garnier added the testing of Grid'5000 tutorials to our continuous integration installation, enabling the earlier detection of problems on the testbed. He then led the migration to PostgreSQL as the backend for the OAR batch scheduler – a behind-the-scenes but major migration.

In addition to daily administrative duties and to his work on Kwapi described below (section 7.3.2), Clément Parisot added support for *production* workloads to Grid'5000, extending the scope of the testbed to make it more suitable for additional user communities. He then managed the installation of the new clusters at Nancy, purchased in the context of OIP Grid'5000 and CPER CyberEntreprises.

Finally, in addition to his roles in the *bureau*, *comité d'architectes* and *comité des responsables de sites* of Grid'5000, Lucas Nussbaum managed the purchase of the new clusters at Nancy mentioned above, and gave several presentations about the testbed, at *Journées SUCCES* [14], at *Retour d'expériences sur la Recherche Reproductible* [15], and at *École Cumulo Numbio*.

7.3.2. A unified monitoring framework for energy consumption and network traffic

Participants: Lucas Nussbaum [contact], Clément Parisot.

Providing experimenters with deep insight about the effects of their experiments is a central feature of testbeds, that Grid'5000 was only partially addressing. We designed Kwapi, a framework that unifies measurements for both energy consumption and network traffic. Because all measurements are taken at the infrastructure level (using sensors in power and network equipment), using this framework has no dependencies on the experiments themselves. Initially designed for OpenStack infrastructures, the Kwapi framework allows monitoring and reporting of energy consumption of distributed platforms. In this work, we extended Kwapi to network monitoring, and overcame several challenges: scaling to a testbed as large as Grid'5000 while still providing high-frequency measurements; providing long-term loss-less storage of measurements; handling operational issues when deploying such a tool on a real infrastructure.

This work was published at Tridentcom [31] and presented in a GENI/FIRE collaboration workshop [12]. It is now in production as the default monitoring framework on Grid'5000.

7.3.3. Comparison of HPC and Clouds testbeds

Participant: Lucas Nussbaum [contact].

Given the recent launch of two large NSF-funded projects that share similar goals as Grid'5000 (CloudLab and ChameleonCloud), we worked on analyzing the design choices made so far by those projects, comparing them with Grid'5000. Preliminary results were presented at REPPAR [17] and at a GENI/FIRE collaboration workshop [13].

7.3.4. Emulation with Distem

Participants: Emmanuel Jeanvoine, Lucas Nussbaum [contact], Cristian Ruiz.

Several improvements have been made around Distem, mostly in the context of ADT COSETTE.

During the internship of Arthur Carcano, we tried to use Distem to experiment on NDN infrastructures. We obtained promising results, especially in terms of scale. We plan to continue this work and publish it in 2016.

We also submitted, to CCGRID, a paper demonstrating the use of Distem to evaluate fault tolerance and load balancing strategies implemented in Charm++. This submission is still pending evaluation.

Finally, in an effort to validate Distem performance, we studied the performance of Container-based virtualization technologies such as LXC or Docker, as most of the underlying technology is also shared with Distem. We studied their performance in the context of HPC, and showed that containers technology has matured over the years, and that performance issues are being solved. This work has been published at VHPC [43].

7.3.5. Management of large-scale experiments

Participants: Emmanuel Jeanvoine, Lucas Nussbaum [contact], Cristian Ruiz.

Following our survey of experiment management tools [7] accepted at FGCS at the end of 2014 and published early this year, we worked on Ruby-Cute, a library that aggregates various useful functionality in the context of such tools. We hope that it will be useful as a basis for future tools, and ease testing of new ideas in that field. The library is available on <http://ruby-cute.github.io/>.

7.3.6. *Tracking provenance in experiment control tools*

Participants: Tomasz Buchert, Lucas Nussbaum [contact].

In the context of our work on XPFlow, we worked on the collection of provenance during experiments. We surveyed provenance collection in various domains of computer science, introduced a new classification of provenance types suited to distributed systems experiments, and proposed a design of a provenance system inspired by this classification. This work has been published at REPPAR [29].

7.3.7. *Reproducible Research*

Participant: Lucas Nussbaum [contact].

Lucas Nussbaum gave a presentation on Reproducible Research[16] at the ICube laboratory seminar (Strasbourg). A shorter version of the talk was given to the Inria *Comité des projets* in Nancy.

Lucas Nussbaum also co-organized the second edition of REPPAR, a workshop on Reproducibility in Parallel Computing, held in conjunction with Euro-Par'2015.

7.4. Routing

7.4.1. *Routing in Wireless Sensor Networks*

Participants: Emmanuel Nataf [contact], Patrick-Olivier Kamgueu, Nesrine Khelifi.

We have formalized our previous work on the routing protocol for wireless sensor network by fuzzy logic specifications. The rules of routing metric composition are now valid for any network depth and we demonstrated its quality by real experimentation [36]. This work is done in the context of the associated team we build with the Cameroun and the Inria international lab LIRIMA.

For potentially very large wireless sensor network, our routing or any other routing, can not limit traffic bottleneck near the network root. Network depth should also be reduced as hop by hop communication is a factor which strongly increases data loss rate. Considering these problems Nesrine Khelifi PhD student of the Manouba University in Tunisia spent 3 months within the Madynes team trying to limit the depth of the network by splitting it under the supervision of network quality observers we had to define.

7.4.2. *Operator calculus based routing in Wireless Sensor Networks*

Participants: Evangelia Tsiontsiou, René Schott, Stacey Staples [Southern Illinois University Edwardsville], Jamilla Benslimane, Bilel Nefzi, Ye-Qiong Song [contact].

Recently, Operator calculus (OC) has been developed by Schott and Staples with whom we collaborate. We make use of OC methods on graphs to solve path selection in the presence of multiple constraints. Based on OC, we developed a distributed algorithm for path selection in a graph. This approach has been applied to efficiently solve a joint routing, channel and time slot scheduling optimization problem in UWB wireless sensor networks [6]. We also designed a new routing protocol which makes use of this algorithm: the Operator Calculus based Routing Protocol (OCRP). In OCRP, a node selects the set of eligible next hops based on the given constraints and the distance to the destination. It then sends the packet to all eligible next hops. The protocol is implemented in Contiki OS (Rime profile) and emulated for TelosB motes using Cooja. We compared its performance against tree and directional flooding routing and showed the advantages of our technique [28]. Our ongoing work consists in its comparison with RPL to show its practical contribution to handle simultaneously several IETF ROLL routing metrics. This work is part of Lorraine AME Satelor project granted by Lorraine Region.

7.4.3. Probabilistic Energy-Aware Routing for Wireless Sensor Networks

Participants: Evangelia Tsiontsiou, Bernardetta Addis, Alberto Ceselli [Universita degli Studi di Milano], Ye-Qiong Song [contact].

Healthcare applications are considered as promising fields for Wireless Sensor Networks (WSNs). Thanks to WSNs, patients can be monitored in hospitals or smart home environments, providing health improvement, or emergency care. A key issue is the limited battery of sensors; indeed, current WSN research trends for healthcare applications include energy efficient routing and network lifetime guarantee mechanisms, among others. One of our ongoing work consists in designing a Smart Probabilistic Energy-Aware Routing Protocol (SPEAR) for WSNs which aims at maximizing the network lifetime by keeping low energy consumption and balancing network traffic between nodes. Our experimental campaign reveals that our SPEAR protocol outperforms the popular Energy Aware Routing Protocol (EAR) from the literature, proving to be more effective in extending the network lifetime. This work has resulted in a conference submission. It is part of Lorraine AME Satelor project granted by Lorraine Region.

7.4.4. Energy-aware IP networks management

Participants: Bernardetta Addis [contact], Giuliana Carello [DEIB, Politecnico di Milano, Italy], Antonio Capone [DEIB, Politecnico di Milano, Italy], Luca Gianoli [Polytechnique de Montreal, Canada], Sara Mattia [IASI, CNR, Roma, Italy], Brunide Sansò [Polytechnique de Montreal, Canada].

The focus of our research is to minimize the energy consumption of the network through a management strategy that selectively switches off devices according to the traffic level. We consider a set of traffic scenarios and jointly optimize their energy consumption assuming a per-flow routing. We propose a traffic engineering mathematical programming formulation based on integer linear programming that includes constraints on the changes of the device states and routing paths to limit the impact on quality of service and the signaling overhead.

A very important issue that may be affected by green networking techniques is resilience to node and link failures. We thus extended the optimization models to guarantee network survivability. Results show that significant savings, up to 30%, may be achieved even when both survivability and robustness are fully guaranteed.

Computational cost of proposed models can be very high when dealing with large size instances (network size and/or number of demands). For this reason, we proposed and tested different problem formulations with the aim of solving larger size instances at optimality. We focus on a particular form of shared protection mechanism, where energy consumption is associated only to active devices during normal functioning. We propose a standard and a projected formulation, with additions of cuts and valid inequalities. Computational results show that the projected formulation is very effective [20]. We plan to extend the work to consider multiperiod scenarios.

7.4.5. Virtual Network Functions Placement and Routing Optimization

Participants: Bernardetta Addis [contact], Dallal Belabed [LIP6, Univ Paris 06, France], Mathieu Bouet [Thales Communications & Security, France], Stefano Secci [LIP6, Univ Paris 06, France].

Network Functions Virtualization (NFV) is incrementally deployed by Internet Service Providers (ISPs) in their carrier networks, by means of Virtual Network Function (VNF) chains, to address customers' demands. The motivation is the increasing manageability, reliability and performance of NFV systems, the gains in energy and space granted by virtualization, at a cost that becomes competitive with respect to legacy physical network function nodes. From a network optimization perspective, the routing of VNF chains across a carrier network implies key novelties making the VNF chain routing problem unique with respect to the state of the art: the bitrate of each demand flow can change along a VNF chain, the VNF processing latency and computing load can be a function of the demands traffic, VNFs can be shared among demands, etc. We started our work providing an integer linear programming model for Virtual Network Functions Placement and demand rerouting. By extensive simulation on realistic ISP topologies, we draw conclusions on the trade-offs achievable between legacy Traffic Engineering (TE) ISP goals and novel combined TE-NFV goals [19].

7.4.6. Composing IoT protocols with Named-Data Networking

Participants: Salvatore Signorello [University of Luxembourg], Olivier Festor [contact], Radu State [University of Luxembourg].

With the emergence of IoT, many layer 2 protocols have been proposed with each of them its own characteristics, advantages and drawbacks. Choosing a protocol often depends on the global context, as for example number of users, time of the day... Although devices can now be fitted with multiple interfaces, using always the same specific layer 2 protocol is not efficient, in particular if we assume that connected devices are retrieving or exchanging similar contents. For example, assuming that WiFi is the most usable interface to download some files in Internet through an access point may not be ideal if a close-by device accessible by Bluetooth already has it. To accommodate so multiple layer 2 protocols, we propose to leverage the Named-Data Networking (NDN) paradigm which allows to explore in parallel multiple paths for retrieving content independently of the underlying protocol. Our first results [46] show that such a theoretical solution cannot work practically. Indeed, applying NDN in a blind mode over multiple layer 2 protocols does not assume the corresponding specificities like for example various collision rates depending on the underlying protocols, which have to be taken into account.

7.5. Multi-modeling and co-simulation

Participants: Laurent Ciarletta [contact], Olivier Festor, Ye-Qiong Song, Yannick Presse, Victorien Elvinger, Julien Vaubourg, Alexandre Tan, Benjamin Segault, Emmanuel Nataf.

Vincent Chevrier (former Maia team, Dep 5, LORIA) is a collaborator and the correspondent for the MS4SG project, Benjamin Camus, and Christine Bourjot (former MAIA team, Dep 5, LORIA) are collaborators for AA4MM/MECSYCO. Julien Vaubourg's PhD is under the co-direction of V. Chevrier and L. Ciarletta.

In Pervasive or Ubiquitous Computing, a growing number of communicating/computing devices are collaborating to provide users with enhanced and ubiquitous services in a seamless way.

These systems, embedded in the fabric of our daily lives, are complex: numerous interconnected and heterogeneous entities are exhibiting a global behavior impossible to forecast by merely observing individual properties. Firstly, users physical interactions and behaviors have to be considered. They are influenced and influence the environment. Secondly, the potential multiplicity and heterogeneity of devices, services, communication protocols, and the constant mobility and reorganization also need to be addressed. Our research on this field is going towards both closing the loop between humans and systems, physical and computing systems, and taming the complexity, using multi-modeling (to combine the best of each domain specific model) and co-simulation (to design, develop and evaluate) as part of a global conceptual and practical toolbox.

We proposed the AA4MM meta-model [76] that solves the core challenges of multimodeling and simulation coupling in an homogeneous perspective. In AA4MM, we chose a multi-agent point of view: a multi-model is a society of models; each model corresponds to an agent and coupling relationships correspond to interaction between agents. In the MS4SG (Multi Simulation for Smart Grids) projet which involves some members of the former MAIA team, Madynes and EDF R&D on smart-grid simulation, we developed a proof of concepts for a smart-appartment case that serves as a basis for building up use cases.

In 2015 we worked on the following research topics:

- Assessment and evaluation of complex systems.
- Cyber Physical Systems

We have led the design and implementation of the Aetournos platform at Loria. The collective movements of a flock of flying communicating robots / UAVs, evolving in potentially perturbed environment constitute a good example of a Cyber Physical System. Applying co-simulation technique we plan to develop a hybrid "network-aware flocking behavior" / "behavior aware routing protocol".

We have provided a working set of tools: multi-simulation behavior / network / physics and generic software development using ROS (Robot Operating System). The UAVs carry a set of sensors for location awareness, their own computing capabilities and several wireless networks.

The effort put in the UAVs gathers academic and research resources from the Aetournos platform, the R2D2 ADT and the 6PO project, while applied, industrial and more R&D projects have been pursued this year (Outback Joe Search and Rescue Challenge, Alerion, Hydradrone) .

- MS4SG to link multi-simulations tools such as HLA (High Level Architecture) and FMI (Functional Mockup Interface) thanks to our AA4MM framework. We have so far successfully applied our solution to the simulation of smart apartment complex and to combine the electrical and networking part of a Smart Grid. The AA4MM software has seen major improvements in 2015 thanks to the resources provided by the MS4SG project and a Carnot engineer financed thanks to Inria. It has been renamed as MECSYCO (<http://www.mecsyco.com>).

Starting from domain specific and heterogeneous models and simulators, the MECSYCO suite allows for multi *systems* integration at several levels: conceptual, formal and software. A couple of visualization tools have been developed as proof of concepts both at run-time and post-mortem.

We have developed software components and plugins that interconnects within MECSYCO heterogeneous simulators from different domains: FMU (working with the 1.0 and 2.0 FMI standard for CoSimulation) or non-FMU such as NS3 or Omnet++.

Several EDF oriented use cases have demonstrated multi-simulations.

In addition to technical reports, several publications have been accepted in 2015 on these subjects [51], [49] and [48].

7.6. Pervasive or Ubiquitous Computing

Participants: Laurent Ciarletta [contact], Olivier Festor, Ye-Qiong Song, Emmanuel Nataf, Thomas Paris, Quentin Houbre, Benjamin Segault, Jonathan Arnault, Eric Perlinski, Antoine Richard.

In Pervasive or Ubiquitous Computing, a growing number of communicating/computing devices are collaborating to provide users with enhanced and ubiquitous services in a seamless way.

These systems, increasingly numerous and heterogeneous, are embedded in the fabric of our daily lives. Our initial subject of interest is to study them with regards to their complexity: Those numerous interconnected and heterogeneous entities are exhibiting a global behavior impossible to forecast by merely observing individual properties.

Firstly, users physical interactions and behaviors have to be considered. They are influenced and influence their surroundings and the environment. Secondly, the potential multiplicity and heterogeneity of devices, services, communication protocols, and the constant mobility and reorganization also need to be addressed.

Our research on this field is going towards both closing the loop between humans and systems, physical and computing systems, and taming the complexity, using multi-modeling (to combine the best of each domain specific model) and co-simulation (to design, develop and evaluate) as part of a global conceptual and practical toolbox.

During some exploratory work, we have seen the potential of these Pervasive Computing resources in the (Very Serious) Gaming area.

In 2015 we worked on the following topics:

- Cyber Physical Systems

We pursued the design and implementation of the Aetournos platform at Loria. The collective movements of a flock of flying communicating robots / UAVs, evolving in potentially perturbed environment constitute a good example of a Cyber Physical System. Eventually, we applied co-simulation technique and plan to develop a hybrid "network-aware flocking behavior" / "behavior aware routing protocol".

We developed a working set of tools: multi-simulation behavior / network / physics and generic software development using ROS (Robot Operating System). The UAVs carry a set of sensor for location awareness, their own computing capabilities and several wireless networks.

The effort put in the UAVs gathers academic and research resources from the Aetournos platform, the Inria ADT R2D2 and the 6PO project, while applied, industrial and more R&D projects have been pursued this year (Medical Express / Outback Joe Search and Rescue Challenge, Alerion, Hydradrone, and a CIFRE PhD with Thales) .

- Smart * (MS4SG)

We have studied scientific problems around model and simulator composition. We have also looked into practical and implementation issues in the frame of our MECSYCO /AA4MM solutions. We have added to our Smart Grid scenarios a smart apartment complex use case.

- (Very Serious) Gaming: Starburst Gaming

7.7. Quality-of-Service

7.7.1. Self-adaptive MAC protocol for both QoS and energy efficiency

Participants: Kévin Roussel, Shuguo Zhuo, Olivier Zendra, Ye-Qiong Song [contact].

Three main contributions have been made this year. Firstly iQueue-MAC has been extended to work on both single channel mode and multi-channel mode, improving its throughput performance [11]. Secondly, S-CoSenS and iQueue-MAC our previously designed protocols have been implemented on RIOT OS over MSP430-based nodes. Our contribution consists in developing a port of RIOT OS on the MSP430 micro-controller and demonstrating that RIOT OS offers rich and advanced real-time features, especially the simultaneous use of as many hardware timers as the underlying platform (micro-controller), which are fundamental features to implement high performance MAC protocols [41]. The Cooja/MSPSim network simulation framework is widely used for developing and debugging, but also for performance evaluation of WSN projects. Our third contribution shows that Cooja is not limited only to the simulation of the Contiki OS based systems and networks, but can also be extended to perform simulation experiments of other OS based platforms, especially that with RIOT OS. Moreover, when performing our own simulations with Cooja and MSPSim, we observed timing inconsistencies with identical experimentations made on actual hardware. Such inaccuracies clearly impair the use of the Cooja/MSPSim framework as a performance evaluation tool, at least for time-related performance parameters. The detailed results of our investigations on the inaccuracy problems, as well as the consequences of this issue, and possible ways to fix or avoid it are available in [42]. Part of this work has been supported by PIA LAR project.

7.7.2. End-to-end delay modeling and evaluation in wireless sensor networks

Participants: François Despaux, Abdelkader Lahmadi, Ye-Qiong Song [contact].

Probabilistic end-to-end performance guarantee may be required when dealing with real-time applications. As part of ANR QUASIMODO project, we are dealing with Markov modeling of multi-hop networks running duty-cycled MAC protocols. One of the problems of the existing Markovian models resides in their strong assumptions that may not be directly used to assess the end-to-end delay in practice. In particular, realistic radio channel, capture effect and OS-related implementation factors are not taken into account. We proposed to explore a new approach combining code instrumentation and Markov chain analysis. In [32] we propose a novel approach to obtain the Markov chain model of sensor nodes by means of Process Mining techniques through the analysis of MAC protocol execution traces for a given traffic scenario. End to end delay is then computed based on this Markov chain. Experimentations were done using IoT-LAB testbed platform. Comparisons in terms of delay have been presented for two different metrics of the RPL protocol (hop count and ETX). The overall approach and its generalization using non-linear regression techniques in terms of traffic rate are detailed in the PhD thesis of François Despaux defended in September 2015 [1].

7.7.3. *Dynamic resource allocation in network virtualization*

Participants: Mohamed Said Seddiki, Mounir Frikha [SupCom, Tunis, Tunisie], Ye-Qiong Song [contact].

This work has been carried out as part of a co-supervised PhD thesis between University of Lorraine and SupCom Tunis.

The objective of this research topic is to develop different resource allocation mechanisms in Network Virtualization, for increasing the QoS guarantee. Firstly, we demonstrated the potential of SDN in the QoS management of a virtualized home network (VN). We proposed and implemented "FlowQoS", a mechanism that can be deployed by an Internet Service Provider in the last-mile hop or in the home gateway. Performance measurements show that this solution can share bandwidth between applications according to user-defined configuration to guarantee QoS for each active traffic. The second contribution is the modeling of the interaction between service providers and infrastructure providers using game theory framework to offer dynamic sharing of physical infrastructure across multiple VN with different QoS requirements. We presented a set of non-cooperative games to model the negotiation phase and the dynamic allocation of nodes and physical links for each deployed VN [10]. Finally we proposed a predictive approach that allows an adaptive control of bandwidth allocation in order to reduce the packet delays for a given VN on each physical link. The last two contributions offer dynamic sharing models of physical infrastructure resources while guaranteeing the QoS for each VN.

The overall approach is detailed in the PhD thesis of Said Seddiki defended in April 2015 [2].

7.7.4. *QoS and fault-tolerance in distributed real-time systems*

Participants: Florian Greff, Laurent Ciarletta, Arnaud Samama [Thales TRT], Eric Dujardin [Thales TRT], Ye-Qiong Song [contact].

The QoS must be guaranteed when dealing with real-time distributed systems interconnected by a network. Not only task schedulability in processors, but also message schedulability in networks should be analysed for validating the system design. Fault-tolerance is another critical issue that one must take into account. In collaboration with Thales TRT industrial partner as part of a CIFRE PhD work, we started a study on the real-time dependability of distributed multi-criticality systems interconnected by an embedded mesh network (RapidIO). For easing the QoS specification at the higher level, DDS middleware is used. We postulate that enhancing QoS for real-time applications entails the development of a cross-layer support of high-level requirements, thus requiring a deep knowledge of the underlying networks. This year, we proposed and implemented a new simulation/emulation/experimentation framework called ERICA, for designing such a feature. ERICA integrates both a network simulator (Ptolemy) and an actual hardware network to allow implementation and evaluation of different QoS-guaranteeing mechanisms. It also supports real-software-in-the-loop, i.e. running of real applications and middleware over these networks. Each component can evolve separately or together in a symbiotic manner, also making teamwork more flexible [68], [33].

7.7.5. *Wireless sensor and actuator networks*

Participants: Lei Mo, Xiufang Shi [Zhejiang University], Jiming Chen [Zhejiang University], Ye-Qiong Song [contact].

Wireless sensor and actuator networks provide a key technology for fully interacting within a CPS (Cyber-Physical System). However, the introduction of the mobile actuator nodes in a network rises some new challenging issues. In this context, we addressed two important issues: the multiple target tracking using both fixed and mobile sensors and the optimal scheduling of mobile wireless energy chargers (actuators) for fixed sensor nodes.

In our work, the data association problem in multiple target tracking is investigated. To reduce the computational complexity of traditional Joint Probabilistic Data Association (JPDA) algorithm, a modified JPDA algorithm is proposed to execute data association in multiple target tracking by utilizing the information of occlusion conditions, which is identified by a three-step algorithm. Simulation results show that the proposed algorithm has good tracking performance but low computational complexity [45].

We also investigated the multiple mobile chargers coordination problem that is minimizing the energy expenditure of the mobile chargers while guaranteeing the perpetual operation of the wireless sensor network. We formulated this problem as a mixed-integer linear program (MILP). To solve this problem efficiently, we proposed a novel decentralized method which is based on Benders decomposition. The multiple mobile chargers coordination problem is then decomposed into a master problem (MP) and a slave problem (SP), with the MP for mobile chargers scheduling and the SP for mobile chargers moving and charging time allocation. The convergence of proposed method is analyzed theoretically. Simulation results demonstrated the effectiveness and scalability of the proposed method [38].

7.7.6. *Big Data-oriented networking*

Participants: Jérôme François [contact], Lautaro Dolberg [University of Luxembourg], Thomas Engel [University of Luxembourg], Raouf Boutaba [University of Waterloo], Reaz Ahmed [University of Waterloo], Shihabur Rahman Chowdhury [University of Waterloo].

Performances of Big Data applications are tightly coupled with the performance of the network in supporting large data transfers. Deploying high-performance networks in data centers is thus vital but configuration and performance management as well as the usage of the network are of paramount importance. We thus surveyed helpful approaches in a book chapter [55]. This chapter starts by discussing the problem of virtual machine placement and its solutions considering the underlying network topology. It then provides an analysis of alternative topologies highlighting their advantages from the perspective of Big Data applications needs. In this context, different routing and flow scheduling algorithms are discussed in terms of their potential for using the network most efficiently. In particular, Software-Defined Networking relying on centralized control and the ability to leverage global knowledge about the network state is propounded as a promising approach for efficient support of Big Data applications.

7.8. **Advanced Cache Management in Content-centric Networks**

Participants: Thomas Silverston [contact], Cholez Thibault, Bernardini César, Aubry Elian, Chrisment Isabelle, Olivier Festor.

Information Centric Networking (ICN) has become a promising new paradigm for the future Internet architecture. It is based on named data, where content address, content retrieval and the content identification is led by its name instead of its physical location. One of the ICN key concepts relies on in-network caching to store multiple copies of data in the network and serve future requests, which helps reducing the load on servers, congestion in the network and enhances end-users delivery performances. Thus, the efficiency of the CCN architecture depends drastically on performances of caching strategies at each node. To date, there has been a lot of studies proposing new caching strategies to improve the performances of CCN. However, among all these strategies, it is still unclear which one performs better as there is a lack of common environment to compare these strategies. To this end, we compared the performances of CCN caching strategies within the same simulation environment. We build a common evaluation scenario and we compare via simulation five relevant caching strategies: Leave Copy Everywhere (LCE), Leave Copy Down (LCD), ProbCache, Cache “Less” For More and MAGIC. We analyze the performances of all the strategies in terms of Cache Hit, Stretch, Diversity and Complexity, and determine the cache strategy that fits the best with every scenario. This work has been published in IEEE Globecom 2015 [26].

At the meantime, CCN architecture uses *Interest* and *Data* messages to request and receive the data, and there has been no routing scheme to match a request to a specific content, as it is currently the case in the Internet. Indeed, CCN relies on flooding, which is a limitation for a future deployment at the Internet-scale. To this end, we proposed a Routing Scheme for CCN based on the softwarization (SDN). In our scheme SRSC, a controller gets knowledge of the network it administers as well as the content, and each node request the next hop to forward the Interest to their controller, until it reaches the closer Content Stores with the requested content. Nodes use a communication channel with the controller that relies only CCN messages and does not use the traditional SDN communication channel protocol Openflow over IP. The rationale is to help having CCN as a stand-alone new networking stack and to enforce its deployment without the IP infrastructure. This research work has been published in IEEE Netsoft 2015 [22] and Algotel 2015 [21].

MAESTRO Project-Team

7. New Results

7.1. Network Science

Participants: Eitan Altman, Konstantin Avrachenkov, Arun Kadavankandy, Jithin Kazhuthuveetil Sreedharan, Hlib Mykhailenko, Philippe Nain, Giovanni Neglia, Yonathan Portilla, Alexandre Reiffers-Masson.

7.1.1. Posting behavior in Social Networks and Content Active Filtering

In [57], Alexandre Reiffers-Masson and Eitan Altman in collaboration with Yezekael Hayel (UAPV), model the posting behavior in Social Networks in topics which have negative externalities, and propose content active filtering in order to increase content diversity. By negative externalities, it is meant that when the quantity of posted contents about some topic increases the popularity of posted contents decreases. They introduce a dynamical model to describe the posting behavior of users taking into account these externalities. Their model is based on stochastic approximations and sufficient conditions are provided to ensure its convergence to a unique rest point. They provide a closed form expression for this rest point. Content Active Filtering (CAF) are actions taken by the administrator of the Social Network in order to promote some objectives related to the quantity of contents posted in various topics. As objective of the CAF they consider maximizing the diversity of posted contents.

7.1.2. Network centrality measures

Recent papers studied the control of spectral centrality measures of a network by manipulating the topology of the network. In [56], Alexandre Reiffers-Masson, Eitan Altman and Yezekael Hayel (UAPV) extend these works by focusing on a specific spectral centrality measure, the Katz-Bonacich centrality. The optimization of the Katz-Bonacich centrality using a topological control is called the Katz-Bonacich optimization problem. The authors first prove that this problem is equivalent to a linear optimization problem. Thus, in the context of large graphs, one can use state-of-the-art algorithms. The authors provide a specific applications of the Katz-Bonacich centrality minimization problem based on the minimization of gossip propagation and make some experiments on real networks which validate the model assumptions.

Betweenness centrality is one of the basic concepts in the analysis of social networks. The initial definition for the betweenness of a node in a graph is based on the fraction of the number of geodesics (shortest paths) between any two nodes that this given node lies on, to the total number of the shortest paths connecting these nodes. This method has quadratic complexity and does not take into account indirect paths. In [45] K. Avrachenkov in collaboration with V. Mazalov (Korelian Institute of Applied Mathematical Research, Russia) and B. Tsynguev (Transbaikal State Univ., Russia) propose a new concept of betweenness centrality for weighted networks, called beta current flow centrality, based on Kirchhoff's law for electric circuits. In comparison with the original current flow centrality and alpha current flow centrality, this new measure can be computed for larger networks. The results of numerical experiments for some examples of networks, in particular, for the popular social network VKontakte as well as the comparison with PageRank method are presented.

PageRank has numerous applications in information retrieval, reputation systems, machine learning, and graph partitioning. In [44], K. Avrachenkov and A. Kadavankandy in collaboration with L.O. Prokhorenkova and A. Raigorodskii (both from Yandex Research) study PageRank in undirected random graphs with expansion property. The Chung-Lu random graph represents an example of such graphs. The authors show that in the limit, as the size of the graph goes to infinity, PageRank can be represented by a mixture of the restart distribution and the vertex degree distribution.

7.1.3. Mining social networks

Social Networks became a major actor in information propagation. Using the Twitter popular platform, mobile users post or relay messages from different locations. The tweet content, meaning and location show how an event-such as the bursty one "JeSuisCharlie" happened in France in January 2015 is comprehended in different countries. In [75], [76] researchers from UAPV and Inria (Mohamed Morchid, Yonathan Portilla, Didier Josselin, Richard Dufour, Eitan Altman, Marc El-Beze, Jean-Valère Cossu, Georges Linarès, Alexandre Reiffers-Masson), studied clustering of the tweets according to the co-occurrence of their terms, including the country, and forecasting the probable country of a non located tweet, knowing its content. First, they present the process of collecting a large quantity of data from the Twitter website. The dataset consists of 2.189 located tweets about "Charlie", from the 7th to the 14th of January. The authors then describe an original method adapted from the Author-Topic (AT) model based on the Latent Dirichlet Allocation method (LDA). They define a homogeneous space containing both lexical content (words) and spatial information (country). During a training process on a part of the sample, the authors provide a set of clusters (topics) based on statistical relations between lexical and spatial terms. During a clustering task, they evaluate the method effectiveness on the rest of the sample that reaches up to 95% of good assignments. It shows that the model is pertinent to foresee tweet location after a learning process.

7.1.4. Analysis of Internet Memes

Memes have been defined by R. Dawkins as cultural phenomena that propagate through non genetic ways. In [42], Eitan Altman and Yonathan Portilla examine three very popular Internet Memes and study their impact on the society in mediterranean countries. the authors use existing software tools (such as Google Trends) as well as tools that they develop in order to quantify the impact of the Memes on the mediterranean societies. The authors obtain quite different results with the different tools they use, which they explain based on some propagation characteristic of each one of the Memes. The analysis shows the extent to which these Memes cross borders and thus contribute to the creation of a globalized culture. The authors finally identify some of the impacts of the globalization of culture.

7.1.5. Trend detection in social networks using Hawkes processes

In [52], Julio Cesar Louzada Pinto and Tijani Chahed (Telecom SudParis) in collaboration with Eitan Altman propose a new trend detection algorithm, designed to find trendy topics being disseminated in a social network. The authors assume that the broadcasts of messages in the social network is governed by a self-exciting point process, namely a Hawkes process, which takes into consideration the real broadcasting times of messages and the interaction between users and topics. The authors formally define trendiness and derive trend indices for each topic being disseminated in the social network. These indices take into consideration the time between the detection and the message broadcasts, the distance between the real broadcast intensity and the maximum expected broadcast intensity, and the social network topology. The proposed trend detection algorithm is simple and uses stochastic control techniques in order to calculate the trend indices. It is also fast and aggregates all the information of the broadcasts into a simple one-dimensional process, thus reducing its complexity and the quantity of data necessary to the detection.

7.1.6. Study of the Youtube recommendation system

The Youtube recommendation system is one the most important view source of a video. In [54], Yonathan Portilla, Alexandre Reiffers-Masson, Eitan Altman in collaboration with Rachid El-Azouzi (UAPV) study the role of recommendation systems in boosting the popularity of videos. The authors first construct a graph that captures the recommendation system in Youtube and study empirically the relationship between the number of views of a video and the average number of views of the videos in its recommendation list. The authors then consider a random walker on the recommendation graph, i.e. a random user that browses through videos such that the video it chooses to watch is selected randomly among the videos in the recommendation list of the previous video it watched. The authors study the stability properties of this random process and show that the trajectory obtained does not contain cycles if the number of videos in the recommendation list is small (which is the case if the computer's screen is small).

7.1.7. Average consensus protocols

In [22] M. El Chamie (Univ. of Washington, USA), G. Neglia and K. Avrachenkov study the weight optimization problem for average consensus protocols by reformulating it as a Schatten norm minimization with parameter p . They show that as p approaches infinity, the optimal solution of the Schatten norm induced problem recovers the optimal solution of the original problem. Moreover, by tuning the parameter p in the proposed minimization, it is possible to trade-off the quality of the solution (i.e., the speed of convergence) for communication/computation requirements (in terms of number of messages exchanged and volume of data processed). They then propose a distributed algorithm to solve the Schatten norm minimization and show that it outperforms the other distributed weight selection methods.

7.1.8. Estimation techniques

The estimation of a large population's size by means of sampling procedures is a key issue in many networking scenarios. Their application domains span from RFID systems to peer-to-peer networks; from traffic analysis to wireless sensor networks; from multicast networks to WLANs. In [14], N. Accettura (Univ. of California Berkeley, USA), G. Neglia and L. A. Grieco (Politecnico di Bari, Italy) illustrate and classify in a coherent framework the main approaches proposed so far in the computer networks literature to deal with such a problem. In particular, starting from the methodologies proposed in ecological studies since the last century, they survey their counterparts in the computer network domain, finding that many lessons can be gained from this insightful investigation. Capture-Recapture techniques are deeply analyzed to allow the reader to exactly understand their pros, cons, and applicability bounds. Finally, they discuss some open issues that deserve further investigations and could be relevant to afford estimation problems in next generation Internet.

Online social networks (OSN) contain extensive amount of information about the underlying society that is yet to be explored. One of the most feasible technique to fetch information from OSN, crawling through Application Programming Interface (API) requests, poses serious concerns over the the guarantees of the estimates. In [70] J. Sreedharan and K. Avrachenkov in collaboration with B. Ribeiro (Purdue University, USA) focus on making reliable statistical inference with limited API crawls. Based on regenerative properties of the random walks, they propose an unbiased estimator for the aggregated sum of functions over edges and proved the connection between variance of the estimator and spectral gap. In order to facilitate Bayesian inference on the true value of the estimator, they derive the approximate posterior distribution of the estimate. Later the proposed ideas are validated with numerical experiments on inference problems in real-world networks.

7.1.9. Percolation in multilayer networks

In [79], P. Nain and his co-authors (S. Guha and P. Basu from Raytheon BB Technologies, D. Towsley from the Univ. of Massachusetts, C. Capar from Ericsson Research, A. Swami from the US Army Research Lab.) consider multiple networks formed by a common set of users connected via M different means of connectivity, where each user (node) is active, independently, in any given network with probability q . They show that when q exceeds a threshold $q_c(M)$, a giant connected component appears in the M -layer network—thereby enabling faraway users to connect using 'bridge' nodes that are active in multiple network layers, even though the individual layers may only have small disconnected islands of connectivity. They show that $q_c(M) \leq \sqrt{\log(1-p_c)}/\sqrt{M}$, where p_c is the bond percolation threshold of the underlying connectivity graph G , and $q_c(1) \equiv q_c$ is its site percolation threshold. The threshold $q_c(M)$ is found explicitly when G is a large random network with an arbitrary node-degree distribution and numerically for various regular lattices. Finally, an intriguingly close connection between this multilayer percolation model and the well-studied problem of site-bond percolation is revealed, in the sense that both models provide a smooth transition between the traditional site and bond percolation models. This connection is used to translate analytical approximations of the site-bond critical region developed in the 1990s, which are functions only of p_c and q_c of the respective lattice, to excellent general approximations of $q_c(M)$.

7.1.10. Extreme Value Theory for Complex Networks

In [20] J. Sreedharan and K. Avrachenkov in collaboration with N. Markovich (Institute of Control Sciences, Moscow) explore the dependence structure in the sampled sequence of complex networks. They consider randomized algorithms to sample the nodes and study extremal properties in any associated stationary sequence of characteristics of interest like node degrees, number of followers, or income of the nodes in online social networks, which satisfy two mixing conditions. Several useful extremes of the sampled sequence like the k th largest value, clusters of exceedances over a threshold, and first hitting time of a large value are investigated. The dependence and the statistics of extremes is abstracted into a single parameter that appears in extreme value theory, called the Extremal Index. The authors derive this parameter analytically and also estimate it empirically. They propose the use of the Extremal Index as a parameter to compare different sampling procedures. As a specific example, degree correlations between neighboring nodes are studied in detail with three prominent random walks as sampling techniques.

7.1.11. Random Matrix Theory for Complex Networks

In [68] A. Kadavankandy and K. Avrachenkov in collaboration with L. Cottatellucci (Eurecom) consider an extension of Erdős-Rényi graph known in the literature as the Stochastic Block Model (SBM). They analyze the limiting empirical distribution of the eigenvalues of the adjacency matrix of a SBM. They derive a fixed point equation for the Stieltjes transform of the limiting eigenvalue empirical distribution function (e.d.f.), concentration results on both the support of the limiting e.d.f. and the extremal eigenvalues outside the support of the limiting e.d.f. Additionally, they derive analogous results for the normalized Laplacian matrix and discuss potential applications of the general results in epidemics and random walks.

In [40], the same authors continue with the analysis of eigenvectors of a Stochastic Block Model. The eigenvalue spectrum of the adjacency matrix of a SBM consists of two parts: a finite discrete set of dominant eigenvalues and a continuous bulk of eigenvalues. They characterize analytically the eigenvectors corresponding to the continuous part: the bulk eigenvectors. For symmetric SBM adjacency matrices, the eigenvectors are shown to satisfy two key properties. A modified spectral function of the eigenvalues, depending on the eigenvectors, converges to the eigenvalue spectrum. Its fluctuations around this limit converge to a Gaussian process different from a Brownian bridge. This latter fact disproves that the bulk eigenvectors are Haar distributed.

7.2. Wireless Networks

Participants: Eitan Altman, Abdulhalim Dandoush.

7.2.1. A General SDN-based IoT Framework with NFV Implementation

The emerging technologies of IoT (Internet of Things), SDN (Software Defined Networking), and NFV (Network Function Virtualization) have a great potential for the information service innovation in the Cloud and big data era. In [26], Jie Li (Tsukuba Univ.) in cooperation with Eitan Altman and with Corinne Touati (Inria Grenoble-Rhône-Alpes) have studied architecture issues in Internet of Things based on SDN with NFV implementation. The contribution of the paper is in providing a view point for integrating these technologies based on their existing standards.

7.2.2. Self-Organizing Network (SON)

Self-Organizing Network (SON) technology aims at autonomously deploying, optimizing and repairing the Radio Access Networks. In [31], Abdoulaye Tall, Zwi Altman (Orange, Issy les Moulineaux) and Eitan Altman showed that in certain cases, it is essential to take into account the impact of the backhaul state in the design of the SON algorithm. They revisit the Base Station load definition taking into account the backhaul state. They provide an analytical formula for the load along with a simple estimator for both elastic and guaranteed bit-rate traffic. They incorporate the proposed load estimator in a self-optimized Load Balancing algorithm. Simulation results for a backhaul constrained heterogeneous network illustrate how the correct load definition can guarantee a proper operation of the SON algorithm.

SON is further studied by these authors in [58], [59] where the Vertical Sectorization (VS) is adapted. VS consists in creating vertically separated sectors in the original cell using an Active Antenna Systems (AAS) supporting two distinct beams with different downtilts. The total transmit power is split between the two sectors, while the frequency bandwidth can be reused by each sector, creating additional interference between the two sectors. For low traffic demand, VS may lead to performance degradation, while for high traffic demand in both sectors, VS is likely to bring about important capacity gains. Hence intelligent activation policy of VS is needed to fully benefit from this feature. The authors propose an approach taking advantage of the more focused downtilted beam. A dynamic alpha fair bandwidth sharing is proposed for low and medium load. It is autonomously replaced by full bandwidth reuse for high load scenarios using a threshold-based controller. A flow-level dynamic simulator is used to numerically validate the proposed mechanisms.

7.2.3. Automated Dynamic Offset for Network Selection in Heterogeneous Networks

Complementing traditional cellular networks with the option of integrated small cells and WiFi access points can be used to further boost the overall traffic capacity and service level. Small cells along with WiFi access points are projected to carry over 60% of all the global data traffic by 2015. With the integration of small cells on the radio access network levels, there is a focus on providing operators with more control over small cell selection while reducing the feedback burden. Altogether, these issues motivate the need for innovative distributed and autonomous association policies that operate on each user under the network operator's control, utilizing only partial information, yet achieving near-optimal solutions for the network. In [25], Majed Haddad (UAPV), Piotr Wiecek (Institute of Mathematics and Computer Science, Wrocław), Saidi Habib (Inria project-team DYOGENE) and Eitan Altman propose a load-aware network selection approach applied to automated dynamic offset in heterogeneous networks. In particular, they investigate the properties of a hierarchical (Stackelberg) Bayesian game framework, in which the macro cell dynamically chooses the offset about the state of the channel in order to guide users to perform intelligent network selection decisions between macro cell and small cell networks. The authors effectively address the problem of how to intelligently configure a dynamic offset which optimizes network's global utility while users maximize their individual utilities.

7.2.4. Localization in ad-hoc wireless sensors networks

Range-based localization algorithms in wireless sensor networks are more accurate but also more computationally complex than the range-free algorithms. The work on this topic by M. S. Elgamel (Arab Academy for Science, Technology & Maritime Transport, Egypt) and A. Dandoush, previously reported, has been published in [23].

7.3. Network Engineering Games

Participants: Eitan Altman, Konstantin Avrachenkov, Giovanni Neglia.

7.3.1. Matching games and the association problem

In [33], Mikaël Touati, Jean-Marc Kélif (Orange Labs), Rachid El-Azouzi (UAPV), Marceau Coupechoux (Telecom ParisTech) and Eitan Altman propose two new algorithms for finding stable structures in ordinal coalition potential games. The first one is enumerative and it performs on a graph. The second one is a modified Deferred Acceptance Algorithm using counter-proposals. It finds a many-to-one matching. The authors illustrate with the example of video caching from a content creator's servers to a service provider's servers.

This is applied to the association of mobiles to IEEE 802.11-based WLANs in populated areas where many mobile terminals are covered by several Access Points (APs) [32]. These mobiles have the possibility to associate to the AP with the strongest signal (best-RSSI association scheme). This can lead to poor performances and overloaded APs. Moreover, the well-known anomaly in the protocol at the MAC layer may also lead to very unpredictable performances and affect the system throughput due to the presence of heterogeneous data rate nodes and the shared nature of the 802.11 medium. In [61], the same authors solve the joint resource allocation and mobile user association after modeling it as a matching game with complementarities, peer effects and selfish players.

7.3.2. Normalized Nash Equilibria for power control with correlated constraints

When correlated constraints are introduced to a game (i.e. the set of actions of a player depends on the policies of other players) there may exist infinitely many Nash equilibria. Assume one wishes to select a particular one u . According to the Karush Kuhn Tucker theorem, there exist Lagrange multipliers such that the best response when all players use their equilibrium policy is the same as that obtained by optimizing the corresponding Lagrangian of that player. The Lagrange multipliers can be interpreted as marginal costs such that if they are imposed on the player as some tax to pay then this induces the player to use Nash equilibrium. The following question arises: does there exist an equilibrium u for which the corresponding Lagrange multipliers are player independent. If the answer is positive then this would make in many cases the billing scalable and simple to implement. An equilibrium u for which the corresponding Lagrange multipliers are player independent is called a normalized Nash equilibrium (NNE). In [39], [50] and [24], Arnob Ghosh (Univ. of Pennsylvania), Laura Cottatellucci (Eurecom) and Eitan Altman provide new conditions for existence and uniqueness of NNE and apply this for power control games arising in cognitive radio [24] and in heterogeneous networks [39], [50].

7.3.3. Admission control to an infinite server queue

In [36], Eitan Altman studies in collaboration with Piotr Wiecek (Wrocław Univ. of Technology) and Arnob Ghosh (Univ. of Pennsylvania) a mean field approximation of the $M/M/\infty$ queueing system. The problem they consider is quite different from standard games of congestion as they consider the case in which higher congestion results in smaller costs per user. This is motivated by a situation in which some TV show is broadcast so that the same cost is needed no matter how many users follow the show. Using a mean-field approximation, they show that this results in multiple equilibria of threshold type which is explicitly computed. The authors further derive the social optimal policy and compute the price of anarchy, and show that the mean-field approximation becomes tight as the workload increases, thus the results obtained for the mean-field model well approximate the discrete one.

7.3.4. Posting Time of Content over a Temporally-Ordered Shared Medium

In [17], Eitan Altman in collaboration with Nahum Shimkin (Technion) consider a game of timing between a random number of content creators, who compete for position and exposure time over a shared medium such as an on-line classified list. Contents (such as ads, messages, multimedia items or comments) are ordered according to their submission times, with more recent submissions displayed at the top (and better) positions. The instantaneous effectiveness of each ad depends on its current display position, as well as on a time-dependent exposure function common to all. Each content creator may choose the submission time of her content within a finite time interval, with the goal of maximizing the total exposure of this content. The authors formulate the problem as a non-cooperative game, analyze its symmetric equilibrium, characterize it in terms of a differential boundary value problem and devise a numerical scheme for its computation.

7.3.5. Routing Games

A central question in routing games has been to establish conditions for the uniqueness of the equilibrium, either in terms of network topology or in terms of costs. This question is well understood in two classes of routing games. The first is the non-atomic routing introduced by Wardrop on 1952 in the context of road traffic in which each player (car) is infinitesimally small; a single car has a negligible impact on the congestion. Each car wishes to minimize its expected delay. Under arbitrary topology, such games are known to have a convex potential and thus a unique equilibrium. The second framework is splittable atomic games: there are finitely many players, each controlling the route of a population of individuals (let them be cars in road traffic or packets in the communication networks). In [64], Eitan Altman and Corinne Touati (Inria Grenoble-Rhône-Alpes) study two other frameworks of routing games in which each of several players has an integer number of connections (which are population of packets) to route and where there is a constraint that a connection cannot be split. Through a particular game with a simple three link topology, they identify various novel and surprising properties of games within these frameworks. The authors show in particular that equilibria are non unique even in the potential game setting of Rosenthal with strictly convex link costs. They further show that non-symmetric equilibria arise in symmetric networks.

7.3.6. Resilience of Routing in Parallel Link Networks

Aniruddha Singhal, Corinne Touati (both from Inria Grenoble-Rhône-Alpes) in collaboration with Eitan Altman and Jie Li (Univ. of Tsukuba) revisit in [63], the resilience problem of routing traffic in a parallel link network model with a malicious player using a game theoretic framework. Consider that there are two players in the network: the first player wishes to split its traffic so as to minimize its average delay, which the second player, i.e., the malicious player, tries to maximize. The first player has a demand constraint on the total traffic it routes. The second player controls the link capacities: it can decrease by some amount the capacity of each link under a constraint on the sum of capacity degradation. The authors first show that the average delay function is convex both in traffic and in capacity degradation over the parallel links and thus does not have a saddle point. They identify best responses strategies of each player and compute both the max-min and the min-max values of the game. One is especially interested in the min-max strategy as it guarantees the best performance under worst possible link capacity degradation. It thus allows to obtain routing strategies that are resilient and robust. The authors compare the results of the min-max to those obtained under the max-min strategies. They provide stable algorithms for computing both max-min and min-max strategies as well as for best responses.

7.3.7. The Social Medium Selection Game

In [72], Fabrice Lebeau (ENS Lyon) Corinne Touati and Nof Abuzainab (Inria Grenoble-Rhône-Alpes) in collaboration with Eitan Altman, consider competition of content creators in routing their content through various media. The routing decisions may correspond to the selection of a social network (e.g. twitter versus facebook or linkedin) or of a group within a given social network. The utility for a player to send its content to some medium is given as the difference between the dissemination utility at this medium and some transmission cost. The authors model this game as a congestion game and compute the pure potential of the game. In contrast to the continuous case, they show that there may be various equilibria. The authors show that the potential is M-concave which allows them to characterize the equilibria and to propose an algorithm for computing it. They then give a learning mechanism which leads to an efficient algorithm to determine an equilibrium. The authors finally determine the asymptotic form of the equilibrium and discuss the implications on the social medium selection problem.

7.3.8. Activation Games in Online Dating Platforms

In [41], Eitan Altman in collaboration with Francesco De Pellegrini (CREATE-NET, Trento) and Huijuan Wang (Delft Univ. of Technology) describe a model for the activation level of users in online dating platforms (ODPs). Such popular systems are conceived in order to match individuals from two groups of potential mates. The business of such platforms pivots around the customers' expectancy to get in contact with their future dates: upon the payment of a fee to the platform owner, ODPs provide specific tools to improve reach and visibility. However, ODPs require a critical number of active users in order to sustain their operations (and their business). Customers of the platform trade off on the price for being more visible and attract mates' contacts. A user becomes inactive if he or she is not contacted by others for some time: being contacted by potential mates acts as an activation signal. The aim of the analysis is to propose a game theoretical framework to capture such a complex activation problem in strategic form. The authors unveil the structure of Nash equilibria and further derive a Stackelberg formulation. The latter is a hierarchical game where the platform owner aims at maximizing profits while preserving the ODP activity level above a critical epidemic threshold.

7.3.9. Epidemics in Networks

Stojan Trajanovski, Huijuan Wang, Piet Van Mieghem (all from Delft Univ. of Technology), in collaboration with Yezekael Hayel (UAPV) and Eitan Altman have pursued their work in the Congas European project concerning malware attacks modeled as SIS (for Susceptible-Infected-Susceptible) epidemics in networks. In [34], the authors consider decentralized optimal protection strategies when a virus is propagating over a network. they assume that each node in the network can fully protect itself from infection at a constant cost, or the node can use recovery software, once it is infected. They model the system using a game theoretic framework and find pure, mixed equilibria, and the Price of Anarchy (PoA) in several network topologies.

Further, they propose both a decentralized algorithm and an iterative procedure to compute a pure equilibrium in the general case of a multiple communities network. Finally, the authors evaluate the algorithms and give numerical illustrations of all results.

They then considered the game-formation problem while balancing multiple, possibly conflicting objectives like cost, high performance, security and resiliency to viruses. In [60], Stojan Trajanovski, Fernando Antonio Kuiper and Piet Van Mieghem (all from Delft Univ. of Technology) in collaboration with Yezekael Hayel (UAPV) and Eitan Altman use a game-formation approach to network design where each player (node), aims to collectively minimize the cost of installing links, of protecting against viruses, and of assuring connectivity. In the game, minimizing virus risk as well as connectivity costs results in sparse graphs. They show that the Nash Equilibria are trees that, according to the Price of Anarchy (PoA), are close to the global optimum, while the worst-case Nash Equilibrium and the global optimum may significantly differ for small infection rate and link installation cost. Moreover, the types of trees, in both the Nash Equilibria and the optimal solution, depend on the virus infection rate, which provides new insights into how viruses spread: for a high infection rate, the path graph is the worst- and the star graph is the best-case Nash Equilibrium. However, for small and intermediate infection rates, trees different from the path and star graphs may be optimal.

7.3.10. Retrial games

In [46] K. Avrachenkov in collaboration with E. Morozov and R. Nekrasova (both from Petrozavodsk State Univ., Russia) consider a single-server retrial system with one and several classes of customers. In the case of several classes, each class has its own orbit for retrying customers. The retrials from the orbits are generated with constant retrial rates. In the single class case, the objective is finding an optimal retrial rate. Whereas in the multi-class case, a game theoretic framework is used and equilibrium retrial rates are found. The performance criteria balance the number of retrials per retrying customer with the number of unhappy customers.

7.3.11. Cooperative Network Design

The Network Design problem has received increasing attention in recent years. Previous works have addressed this problem considering almost exclusively networks designed by selfish users, which can be consistently suboptimal. In [18] K. Avrachenkov, J. Elias (Univ. Paris Descartes, France), F. Martignon (Univ. Paris Sud, France), G. Neglia and L. Petrosyan (St. Petersburg State Univ.) address the network design issue using cooperative game theory, which permits to study ways to enforce and sustain cooperation among users. Both the Nash bargaining solution and the Shapley value are widely applicable concepts for solving these games. However, the Shapley value presents several drawbacks in this context. For this reason, they solve the cooperative network design game using the Nash bargaining solution (NBS) concept. More specifically, they extend the NBS approach to the case of multiple players and give an explicit expression for users' cost allocations. They further provide a distributed algorithm for computing the Nash bargaining solution. Then, they compare the NBS to the Shapley value and the Nash equilibrium solution in several network scenarios, including real ISP topologies, showing its advantages and appealing properties in terms of cost allocation to users and computation time to obtain the solution.

Numerical results demonstrate that the proposed Nash bargaining solution approach permits to allocate costs fairly to users in a reasonable computation time, thus representing a very effective framework for the design of efficient and stable networks.

7.4. Green Networking and Smart Grids

Participants: Sara Alouf, Eitan Altman, Alberto Benegiamo, Alain Jean-Marie, Giovanni Neglia.

7.4.1. Energy efficiency and management in wireless networks

In [35], Rodrigo A. Vaca Ramirez and John S. Thompson (Univ. of New England), in collaboration with Eitan Altman and Victor Ramos Ramos (UAM - Univ. Autonoma Metropolitana Unidad Iztapalapa) consider a low complexity virtual Multiple-input Multiple-output (MIMO) coalition formation algorithm. The goal is to obtain improvements in energy efficiency by forming multi-antenna virtual arrays for information

transmission in the uplink. Virtual arrays are formed by finding a stable match between single antenna devices such as mobile station (MS) and relay stations (RS) by using a game theoretic approach derived from the concept of the college admissions problem. They focus on enhancing the MS performance by forming virtual coalitions with the RSs. Thus, power savings are obtained through multi-antenna arrays by implementing the concepts of spatial diversity and spatial multiplexing for uplink transmission. They focus on optimizing the overall consumed power rather than the transmitted power of the network devices. Furthermore, it is shown analytically and by simulations that when overall consumed power is considered as an optimization metric, the energy efficiency of the single antennas devices is not always improved by forming a virtual MIMO array. Hence, single antenna devices may prefer to transmit on their own when channel conditions are favorable. In addition, the simulation results show that the proposed framework provides comparable energy savings and a lower implementation complexity when compared to a centralized exhaustive search approach that is coordinated from the Base Station.

Sara Alouf, Ioannis Dimitriou (now at Univ. Patras, Greece) and Alain Jean-Marie had worked on the modeling of wireless communication base stations with autonomous energy supply (solar, wind). They had proposed a versatile 5-dimensional Markov model of the device, and shown that the Quasi Birth-Death framework is adequate for solving the model. This work has been completed with a companion product-form model based on E. Gelenbe's modeling of energy networks with signals [48].

7.4.2. Stochastic Geometric Models for Green Networking

In [16], Eitan Altman in collaboration with Cengiz Hasan, Manjesh Kumar Hanawal (IIT Mumbai), Shlomo Shamai (Technion), Jean-Marie Gorce (Inria project-team SOCRATE), Rachid El-Azouzi (UAPV) and Laurent Roullet (Alcatel Lucent Bell Labs) study both the uplink and downlink energy efficiency based on the assumption that base stations are distributed according to an independent stationary Poisson point process. This type of modeling allows to make use of the property that the spatial distribution of the base stations after thinning (switching-off) is still a Poisson process. This implies that the probability of the SINR can be kept unchanged when switching-off base stations provided that one scales up the transmission power of the remaining base stations. The authors then solve the problem of optimally selecting the switch-off probabilities so as to minimize the energy consumptions while keeping unchanged the SINR probability distribution. They then study the trade-off in the uplink performance involved in switching-off base stations. These include energy consumption, the coverage and capacity, and the impact on amount of radiation absorbed by the transmitting user.

7.4.3. Direct Load Control

Energy demand and production need to be constantly matched in the power grid. The traditional paradigm to continuously adapt the production to the demand is challenged by the increasing penetration of more variable and less predictable energy sources, like solar photovoltaics and wind power. An alternative approach is the so called direct control of some inherently flexible electric loads to shape the demand. Direct control of deferrable loads presents analogies with flow admission control in telecommunication networks: a request for network resources (bandwidth or energy) can be delayed on the basis of the current network status in order to guarantee some performance metrics. In [53] G. Neglia, in collaboration with G. Di Bella (Telecom Italia, Italy), L. Giarré and I. Tinnirello (Univ. of Palermo, Italy) go beyond such an analogy, showing that usual teletraffic tools can be effectively used to control energy loads. In particular they propose a family of control schemes which can be easily tuned to achieve the desired trade-off among resources usage, control overhead and privacy leakage.

7.4.4. Charge of Electric Vehicles

The massive introduction of Electric Vehicles (EVs) will make fleet managers spend a significant amount of money to buy electric energy. If energy price changes over time, accurate scheduling of recharging times may result in significant savings. In [29] C. Rottondi (IDSIA Dalle Molle Institute for Artificial Intelligence, Switzerland), G. Neglia and G. Verticale (Politecnico di Milano, Italy) evaluate the complexity of the optimal scheduling problem considering a scenario with a fleet manager having full knowledge of the customers'

traveling needs at the beginning of the scheduling horizon. They prove that the problem has polynomial complexity and provide complexity lower and upper bounds. Moreover, they propose an online sub-optimal scheduling heuristic that schedules the EVs' recharge based on historical travelling data. They compare the performance of the optimal and sub-optimal methods to a benchmark online approach that does not rely on any prior knowledge of the customers' requests, in order to evaluate whether the additional complexity required by the proposed strategies is worth the achieved economic advantages. Numerical results show up to of 35% cost savings with respect to the benchmark approach.

7.5. Content-Oriented Systems

Participants: Sara Alouf, Eitan Altman, Konstantin Avrachenkov, Alain Jean-Marie, Philippe Nain, Giovanni Neglia.

7.5.1. Modeling modern DNS caches

Sara Alouf and Nicaise Choungmo Fofack (former PhD student at MAESTRO, currently at Ingima) have thoroughly revised their study of the modern behavior of DNS caches. In particular the closure properties of the class of distributions called *diagonal matrix-exponential* are fully derived, hence the analytic models presented in [78] to study tree of caches with general caching durations are extended to the case of polytrees [15].

7.5.2. Data placement and retrieval in distributed/peer-to-peer systems

In previous years, Alain Jean-Marie and collaborators from the Univ. Montpellier have defined a family of combinatorial designs that minimize the variance in the availability of replicated documents in unreliable infrastructures. Then with Jean-Claude Bermond (CNRS, with the Inria project-team COATI), Dorian Mazauric (now with Inria project-team ABS) and Joseph Yu (UFV Vancouver), it was shown that *well-balanced families* solve the problem, and such families were constructed for small numbers of replicas. This work is now published in [21]. During the internship of Mikhail Grigorev, several methods for generating at random good solutions have been investigated.

7.5.3. Fairness in caching systems

Data offloading from the cellular network to lowcost WiFi has been the subject of several research works in the last years. In-network caching has also been studied as an efficient means to further reduce cellular network traffic. In [49] M. El Chamie (Univ. of Washington, USA), C. Barakat (Inria project-team DIANA) and G. Neglia consider a scenario where mobile users can download popular contents (e.g., maps of a city, shopping information, social media, etc.) from WiFi-enabled caches deployed in an urban area. They study the optimal distribution of contents among the caches (i.e., what contents to put in each cache) to minimize users' access cost in the whole network, and argue that this optimal distribution does not necessarily provide geographic fairness, i.e., users at different locations can experience highly variable performance. In order to mitigate this problem, they propose two different cache coordination algorithms based on gossiping. These algorithms achieve geographic fairness while preserving the minimum access cost for end users.

In [43] K. Avrachenkov in collaboration with V.S. Borkar (IIT Mumbai, India) consider the task of scheduling a crawler to retrieve from several sites their ephemeral content. This is content, such as news or posts at social network groups, for which a user typically loses interest after some days or hours. Thus development of a timely crawling policy for ephemeral information sources is very important. The authors first formulate this problem as an optimal control problem with average reward. The reward can be measured in terms of the number of clicks or relevant search requests. The problem in its exact formulation suffers from the curse of dimensionality and quickly becomes intractable even with moderate number of information sources. Fortunately, this problem admits a Whittle index, a celebrated heuristics which leads to problem decomposition and to a very simple and efficient crawling policy. The Whittle index is derived, together with its theoretical justification.

7.6. Advances in Methodological Tools

Participants: Eitan Altman, Konstantin Avrachenkov, Ilaria Brunetti.

7.6.1. Control theory

In [19] K. Avrachenkov in collaboration with O. Habachi (UAPV) and A. Piunovskiy and Y. Zhang (both from Univ. of Liverpool, UK) investigate infinite horizon deterministic optimal control problems with both gradual and impulsive controls, where any finitely many impulses are allowed simultaneously. Both discounted and long run time average criteria are considered. They establish very general and at the same time natural conditions, under which the dynamic programming approach results in an optimal feedback policy. The established theoretical results are applied to the Internet congestion control, and by solving analytically and nontrivially the underlying optimal control problems, they obtain a simple threshold-based active queue management scheme, which takes into account the main parameters of the transmission control protocols, and improves the fairness among the connections in a given network.

7.6.2. Game theory

7.6.2.1. Evolutionary games

Standard Evolutionary Game framework is a useful tool to study large interacting systems and to understand the strategic behavior of individuals in such complex systems. Adding an individual state to model a local feature of each player in this context, allows one to study a wider range of problems in various application areas as networking, biology, etc. In [47], Ilaria Brunetti and Eitan Altman in collaboration with Yezekael Hayel (UAPV) introduce such an extension of evolutionary game framework and particularly, focus on the dynamical aspects of this system. Precisely, the authors study the coupled dynamics of the strategies and the individual states inside a population of interacting individuals. They consider here a two-strategies evolutionary game. They first obtain a system of combined dynamics and they show that the rest-points of this system are equilibria of the evolutionary game with individual state. Second, by assuming two different time scales between states and strategy dynamics, one can compute explicitly the equilibria. Then, by transforming the evolutionary game with individual states into a standard evolutionary game, the authors obtains an equilibrium which is equivalent, in terms of occupation measure, to the previous one. Finally, they show a generalization of the model. All the results are illustrated with numerical results.

7.6.2.2. Stochastic Games

Motivated by uncertainty in the value of the interest rate, in [62] K. Avrachenkov in collaboration with A. Varava (KTH, Sweden) study discounted zero-sum stochastic games with an arbitrary discount factor. Their general goal is to obtain a power series expansion of the value of the game with respect to the discount factor around its nominal value. They consider a specific but important class of stochastic games – completely mixed stochastic games. As an illustrative example they take a tax evasion model.

MESCAL Project-Team

6. New Results

6.1. Reproducible Research

In the field of large-scale distributed systems, experimentation is particularly difficult. The studied systems are complex, often non-deterministic and unreliable, software is plagued with bugs, whereas the experiment workflows are unclear and hard to reproduce. In [11], we provide an extensive list of features offered by general-purpose experiment management tools dedicated to distributed systems research on real platforms. We then use it to assess existing solutions and compare them, outlining possible future paths for improvements.

In [20], we address the question of developing a lightweight and effective workflow for conducting experimental research on modern parallel computer systems in a reproducible way. Our workflow simply builds on two well-known tools (Org-mode and Git) and enables us to address issues such as provenance tracking, experimental setup reconstruction, replicable analysis. Although this workflow is perfectible and cannot be seen as a final solution, we have been using git for two years now and we have recently published a fully reproducible article, which demonstrates the effectiveness of our proposal.

6.2. Performance Characterization and Optimization of IOs

In high-performance computing environments, parallel file systems provide a shared storage infrastructure to applications. In the situation where multiple applications access this shared infrastructure concurrently, their performance can be impaired because of interference. In [22], we improve performance by alleviating interference effects through a smart I/O scheduler that organizes and optimizes the applications' requests and adjusts the access pattern to the device characteristics. We apply machine learning techniques to automatically select the best scheduling algorithm for each situation. Our approach improves performance by up to 75

In [33], we present a new storage device profiling tool that characterizes the sequential to random throughput ratio for reads and writes of different sizes. As we explained previously, several optimizations aim at adapting applications' access patterns in order to generate contiguous accesses for improved performance when accessing storage devices like hard disks. However, when considering other storage options like RAID arrays and SSDs, the access time ratio between contiguous and non-contiguous accesses may not compensate for these optimizations' cost. In this scenario, the information provided by our tool could be used to dynamically decide if optimizations are beneficial for performance, which is why we took a particular attention to obtain accurate information in a minimal benchmarking time.

6.3. Application of Game Theory and Distributed Optimization to Wireless Networks

In wireless networks, channel conditions and user quality of service (QoS) requirements vary, often quite arbitrarily, with time (e.g. due to user mobility, fading, etc.) and users only have a very limited information about their environment. In such context optimizing transmission while taking power consumption into account is extremely challenging. We apply game theory technique to MIMO wireless network using OFDM or OFDMA where multi-path channels can be handled efficiently

In [25], [9], we show that distributed power allocation in heterogeneous OFDMA cognitive radio networks can be modeled as a game where each user equipment in the network engages in a non-cooperative game and allocates its available transmit power over subcarriers to maximize its individual utility. The corresponding equilibrium (Debreu, an extension of Nash Equilibrium) can be characterized with fractional programming and we provide sufficient conditions for computing such equilibria as fixed points of a water-filling best response operator. Using such approach can however be quite slow and is very sensitive to delay and information uncertainty (it may not converge). Therefore, we explain in [17] how signal covariance matrices in Gaussian MIMO multiple access channel can be learnt in presence of imperfect (and possibly delayed) feedback. The algorithm we propose is based on the method of matrix exponential learning (MXL) and it has the same information and computation requirements as distributed water-filling. However our algorithm converge much faster even for large numbers of users and/or antennas per user and in the presence of user update asynchrony, random delays and/or ergodically changing channel conditions. Yet, since the system may evolve over time in an unpredictable fashion (e.g. due to changes in the wireless medium or the users' QoS requirements), static solution concepts (such as Nash equilibrium) may be no longer relevant and users must adapt to changes in the environment "on the fly", without being able to predict the system's evolution ahead of time. Hence, we focus on the concept of no-regret : policies that perform at least as well as the best fixed transmit profile in hindsight. In [31] and [41], we provide a formulation of power control as an online optimization problem and we show that the FM dynamics lead to no regret in this dynamic context. In [40] we apply this approach energy efficient transmission in MIMO-OFDM systems and we show through numerical simulations that, in realistic network environments even under rapidly changing channel conditions, users can track their individually optimum transmit profile, achieving gains of up to 600 in energy efficiency over uniform power allocation policies.

We also apply this technique to multi-carrier cognitive radio systems. Such systems allow opportunistic secondary users (SUs) to access portions of the spectrum that are unused by the network's licensed primary users (PUs), provided that the induced interference does not compromise the PUs' performance guarantees. In [14], we introduce a flexible spectrum access pricing schemes such that the corresponding Nash equilibrium is unique under very mild assumptions and satisfies the performance constraints. In addition, we derive a dynamic power allocation policy that converges to equilibrium within a few iterations (even for large numbers of users) and that relies only on local—and possibly imperfect—signal-to-interference-and-noise ratio measurements. In [24], we draw on exponential learning techniques to design an algorithm that is able to adapt to system changes "on the fly", i.e. such that the proposed transmit policy leads to no regret even under rapidly changing network conditions.

6.4. General Results in Game Theory

Our work on game theory is often motivated by applications to wireless networks but can often have a more general application.

In [38], motivated by applications to multi-antenna wireless networks, we propose a distributed and asynchronous algorithm for stochastic semidefinite programming. This algorithm is a stochastic approximation of a continuous-time matrix exponential scheme regularized by the addition of an entropy-like term to the problem's objective function. We show that the resulting algorithm converges almost surely to an (ϵ) -approximation of the optimal solution requiring only an unbiased estimate of the gradient of the problem's stochastic objective.

As explained in the previous section, classical Nash equilibrium concepts become irrelevant in situations where the environment evolves over time. In [15], we study one of the main concept of online learning and sequential decision problem known as regret minimization. Our objective is to provide a quick overview and a comprehensive introduction to online learning and game theory.

In practice, it is rarely reasonable to assume that players have access to the strategy of the others and implementing a best response can thus become cumbersome. Replicator dynamics is a fundamental approach in evolutionary game theory in which players adjust their strategies based on their actions' cumulative payoffs over time – specifically, by playing mixed strategies that maximize their expected cumulative payoff.

- In [19], we investigate the impact of payoff shocks on the evolution of large populations of myopic players that employ simple strategy revision protocols such as the "imitation of success". In the

noiseless case, this process is governed by the standard (deterministic) replicator dynamics; in the presence of noise however, the induced stochastic dynamics are different from previous versions of the stochastic replicator dynamics (such as the aggregate-shocks model of Fudenberg and Harris, 1992). In this context, we show that strict equilibria are always stochastically asymptotically stable, irrespective of the magnitude of the shocks; on the other hand, in the high-noise regime, non-equilibrium states may also become stochastically asymptotically stable and dominated strategies may survive in perpetuity (they become extinct if the noise is low). Such behavior is eliminated if players are less myopic and revise their strategies based on their cumulative payoffs. In this case, we obtain a second order stochastic dynamical system whose attracting states coincide with the game's strict equilibria and where dominated strategies become extinct (a.s.), no matter the noise level.

- In [13], we study a new class of continuous-time learning dynamics consisting of a replicator-like drift adjusted by a penalty term that renders the boundary of the game's strategy space repelling. These penalty-regulated dynamics are equivalent to players keeping an exponentially discounted aggregate of their ongoing payoffs and then using a smooth best response to pick an action based on these performance scores. Building on the duality with evolutionary game theory, we design a discrete-time, payoff-based learning algorithm that converges to (arbitrarily precise) approximations of Nash equilibria in potential games. Moreover, the algorithm remains robust in the presence of stochastic perturbations and observation errors, and it does not require any synchronization between players, which is a very important property when applying such technique to traffic engineering.
- In [18], we investigate an other class of reinforcement learning dynamics in which the players strategy adjustment is regularized with a strongly convex penalty term. In contrast to the class of penalty functions used to define smooth best responses in models of stochastic fictitious play, the regularizers used in this paper need not be infinitely steep at the boundary of the simplex. Dropping this requirement gives rise to an important dichotomy between steep and non-steep cases. In this general setting, our main results extend several properties of the replicator dynamics such as the elimination of dominated strategies, the asymptotic stability of strict Nash equilibria and the convergence of time-averaged trajectories to interior Nash equilibria in zero-sum games.
- In [37], we study a general class of game-theoretic learning dynamics in the presence of random payoff disturbances and observation noise, and we provide a unified framework that extends several rationality properties of the (stochastic) replicator dynamics and other game dynamics. In the unilateral case, we show that the stochastic dynamics under study lead to no regret, irrespective of the noise level. In the multi-player case, we find that dominated strategies become extinct (a.s.) and strict Nash equilibria remain stochastically asymptotically stable – again, independently of the perturbations' magnitude. Finally, we establish an averaging principle for 2-player games and we show that the empirical distribution of play converges to Nash equilibrium in zero-sum games under any noise level.

6.5. Simulation

Simgrid is a toolkit providing core functionalities for the simulation of distributed applications in heterogeneous distributed environments. Although it was initially designed to study large distributed computing environments such as grids, we have recently applied it to performance prediction of HPC configurations.

- Indeed, multi-core architectures comprising several GPUs have become mainstream but obtaining the maximum performance of such heterogeneous machines is challenging as it requires to carefully offload computations and manage data movements between the different processing units. The most promising and successful approaches so far build on task-based runtimes that abstract the machine and rely on opportunistic scheduling algorithms. As a consequence, the problem gets shifted to choosing the task granularity, task graph structure, and optimizing the scheduling strategies. Trying different combinations of these different alternatives is also itself a challenge. Indeed, getting accurate measurements requires reserving the target system for the whole duration of experiments. Furthermore, observations are limited to the few available systems at hand and may be difficult

to generalize. In [21], we show how we crafted a coarse-grain hybrid simulation/emulation of StarPU, a dynamic runtime for hybrid architectures, over SimGrid. This approach allows to obtain performance predictions of classical dense linear algebra kernels accurate within a few percents and in a matter of seconds, which allows both runtime and application designers to quickly decide which optimization to enable or whether it is worth investing in higher-end GPUs or not. Additionally, it allows to conduct robust and extensive scheduling studies in a controlled environment whose characteristics are very close to real platforms while having reproducible behavior. In [30], we have extended this approach to the simulation of a multithreaded multifrontal QR solver of sparse matrices: QR-MUMPS. In our approach, the target high-end machines are calibrated only once to derive sound performance models. These models can then be used at will to quickly predict and study in a reproducible way the performance of such irregular and resource-demanding applications using solely a commodity laptop. Our approach also allows to study the memory consumption along time, which is a critical factor for such applications.

- Beside the inherent heterogeneity of distributed computing infrastructures, storage is also an essential component to cope with the tremendous increase in scientific data production and the ever-growing need for data analysis and preservation. Understanding the performance of a storage subsystem or dimensioning it properly is an important concern for which simulation can help. In [29], we detail how we have extended SimGrid with storage simulation capacities and we list several concrete use cases of storage simulations in clusters, grids, clouds, and data centers for which the proposed extension would be beneficial.

Ψ^2 is a simulation software of markovian models that is able to provide a perfect sampling of the stationary distribution. In [12], we consider open Jackson networks with losses with mixed finite and infinite queues and analyze the efficiency of sampling from their exact stationary distribution. We show that perfect sampling is possible, although the underlying Markov chain may have an infinite state space. The main idea is to use a Jackson network with infinite buffers (that has a product form stationary distribution) to bound the number of initial conditions to be considered in the coupling from the past scheme. We also provide bounds on the sampling time of this new perfect sampling algorithm for acyclic or hyper-stable networks. These bounds show that the new algorithm is considerably more efficient than existing perfect samplers even in the case where all queues are finite. We illustrate this efficiency through numerical experiments. We also extend our approach to variable service times and non-monotone networks such as queueing networks with negative customers.

6.6. Asymptotic Models

Analyzing a set of n stochastic entities interacting with each others can be particularly difficult but the *mean field approximation* is a very effective technique to characterize the probability distribution of such systems when the number of entities n grows very large. The limit system is generally deterministic and characterized by a differential equation that is more amenable to analysis and optimization. Such approximation however typically requires that the dynamics of the entities depend only on their state (the state space of each object does not scale with n the number of objects) but neither on their identity nor on their spatial location.

- In [28], we analyze a family of list-based cache replacement algorithms. We present explicit expressions for the cache content distribution and miss probability under some assumptions and we develop an algorithm with a time complexity that is polynomial in the cache size and linear in the number of items to compute the exact miss probability. We further introduce a mean field model to approximate the transient behavior of the miss probability and prove that this model becomes exact as the cache size and number of items tends to infinity. We show that the set of ODEs associated to the mean field model has a unique fixed point that can be used to approximate the miss probability in case the exact computation becomes too time consuming. Using this approximation, we provide guidelines on how to select a replacement algorithm within the family considered such that a good trade-off is achieved between the cache reactivity and its steady-state hit probability
- For distributed systems where /locality/ is essential in the dynamics the mean-field approach requires to resort to discretization of space into a finite number of cells to fit in the classical framework.

Such approach not only scales badly but also requires that spatial interactions are weak. One of the tool to tackle this difficult problem comes from statistical physics and is popular in biology: pair approximation. In [26], we successfully apply this approach to the "Power of Two Choice" load balancing paradigm: each incoming task is allocated to the least loaded of two servers picked at random among a collection of n servers. We study the power of two-choice in a setting where the two servers are not picked independently at random but are connected by an edge in an underlying graph. Our problem is motivated by systems in which choices are geometrically constrained (e.g., a bike-sharing system). We study a dynamic setting in which jobs leave the system after being served by a server to which it was allocated. Our focus is when each server has few neighbors (typically 2 to 4) for which a mean-field approximation is not accurate. We build the pair-approximation equations and show that they describe accurately the steady-state of the system. Our results show that, even in a graph of degree 2, choosing between two neighboring servers improve dramatically the performance compared to a random allocation.

- In [8], we consider a queueing system composed of a dispatcher that routes deterministically jobs to a set of non-observable queues working in parallel. In this setting, the fundamental problem is which policy should the dispatcher implement to minimize the stationary mean waiting time of the incoming jobs. We present a structural property that holds in the classic scaling of the system where the network demand (arrival rate of jobs) grows proportionally with the number of queues. Assuming that each queue of type r is replicated k times, we consider a set of policies that are periodic with period $k \sum_r p_r$ and such that exactly p_r jobs are sent in a period to each queue of type r . When $k \rightarrow \infty$, our main result shows that all the policies in this set are equivalent, in the sense that they yield the same mean stationary waiting time, and optimal, in the sense that no other policy having the same aggregate arrival rate to all queues of a given type can do better in minimizing the stationary mean waiting time. Furthermore, the limiting mean waiting time achieved by our policies is a convex function of the arrival rate in each queue, which facilitates the development of a further optimization aimed at solving the fundamental problem above for large systems.

6.7. Trace and Statistical Analysis

Although we often use Markovian approaches to model large scale distributed system, these probabilistic tools can also be used to lay the foundation of statistical analysis of traces of real systems.

- In [36], we explain how we apply statistical modelling and statistical inference of the ANR GEOMEDIA corpus, that is a collection of international RSS news feeds. Central to this project, RSS news feeds are viewed as a representation of the information in geopolitical space. As such they allow us to study media events of global extent and how they affect international relations. Here we propose hidden Markov models (HMM) as an adequate modelling framework to study the evolution of media events in time. This set of models respect the characteristic properties of the data, such as temporal dependencies and correlations between feeds. Its specific structure corresponds well to our conceptualisation of media attention and media events. We specify the general model structure that we use for modelling an ensemble of RSS news feeds. Finally, we apply the proposed models to a case study dedicated to the analysis of the media attention for the Ebola epidemic which spread through West Africa in 2014.
- The use of stochastic formalisms, such as Stochastic Automata Networks (SAN), can be very useful for statistical prediction and behavior analysis. Once well fitted, such formalisms can generate probabilities about a target reality. These probabilities can be seen as a statistical approach of knowledge discovery. However, the building process of models for real world problems is time consuming even for experienced modelers. Furthermore, it is often necessary to be a domain specialist to create a model. In [34], we present a new method to automatically learn simple SAN models directly from a data source. This method is encapsulated in a tool called SAN Generator (SANGE). Through examples we show how this new model fitting method is powerful and relatively easy to use, which can grant access to a much broader community to such powerful modeling formalisms.

- In [32], we have presented our recent results on macroscopic analysis of huge traces of parallel/distributed applications. To identify a *macroscopic phenomenon* over large traces, one needs to change the representation scale and to aggregate data both in time, space and application structure through meaningful operators to propose *multi-scale visualizations*. The question is then to know the quantity of information lost by such scaling to be able to correctly interpret them. The principles underlying this approach are based on information theory since the conditional entropy of an aggregation indicates the quantity of information loss when data are aggregated. This approach has been integrated in the Framesoc framework [35].
- In [27], We study the problem of making forecasts about the future availability of bicycles in stations of a bike-sharing system (BSS). This is relevant in order to make recommendations guaranteeing that the probability that a user will be able to make a journey is sufficiently high. To this end, we use probabilistic predictions obtained from a queuing theoretical time-inhomogeneous model of a BSS. The model is parametrized and successfully validated using historical data from the Vélib ' BSS of the City of Paris. We develop a critique of the standard root-mean-square-error (RMSE), commonly adopted in the bike-sharing research as an index of the prediction accuracy, because it does not account for the stochasticity inherent in the real system. Instead we introduce a new metric based on scoring rules. We evaluate the average score of our model against classical predictors used in the literature. We show that these are outperformed by our model for prediction horizons of up to a few hours. We also discuss that, in general, measuring the current number of available bikes is only relevant for prediction horizons of up to few hours.

MIMOVE Team

7. New Results

7.1. Introduction

MiMove's research activities in 2015 have focused on a set of areas directly related to the team's research topics. Hence, we have worked on QoS for Emergent Mobile Systems (§ 7.2) in relation to our research topic regarding Emergent Mobile Distributed Systems (§ 3.2). Furthermore, our effort on SoundCity (§ 7.3) is linked to our research on Mobile Social Crowd-sensing (§ 3.4). Still in the context of Mobile Social Crowd-sensing (§ 3.4), we have developed AppCivist-PB (§ 7.4) related to our interest in social applications aiming to actively involve citizens (see § 4.1); this is further linked to our research on composition of Emergent Mobile Distributed Systems (§ 3.2).

7.2. QoS for Emergent Mobile Systems

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Rachit Agarwal, Valérie Issarny, Raphael de Aquino Gomes.

With the emergence of Future Internet applications that connect web services, sensor-actuator networks and service feeds into open, dynamic, mobile choreographies, heterogeneity support of interaction paradigms is of critical importance. Heterogeneous interactions can be abstractly represented by client-server, publish/subscribe, tuple space and data streaming middleware connectors that are interconnected via bridging mechanisms providing interoperability among the choreography peers. We make use of the *eVolution Service Bus (VSB)* (see § 6.2) as the connector enabling interoperability among heterogeneous choreography participants. VSB models interactions among peers through generic *post* and *get* operations that represent peer behavior with varying time/space coupling.

Within this context, we study end-to-end Quality of Service (QoS) properties of choreographies, where in particular we focus on the effect of middleware interactions on QoS. We consider both homogeneous and heterogeneous (via VSB) interactions. We report in the following our results in three complementary directions:

- While VSB ensures functional interoperability of heterogeneous choreography interactions, differences in timing requirements and constraints of such interactions can severely affect their latencies and success rates. To model timeliness, we introduce the *lease* and *timeout* parameters. The former captures data availability and validity in time, while the latter represents intermittent availability of data recipients due to mobility and disconnection. By precisely studying the related timing thresholds using timed automata models, we verify conditions for successful interactions with VSB connectors. Furthermore, we statistically analyze through simulations, the effect of varying lease and timeout periods to ensure higher probabilities of successful interactions. Simulation experiments are compared with experiments run on the VSB implementation testbed to evaluate the accuracy of results. This work can provide application developers with precise design time information when setting these timing thresholds in order to ensure accurate runtime behavior [23].
- Choreography peers deployed in mobile environments are typically characterized by intermittent connectivity and asynchronous reception of data. In such environments, it is essential to guarantee acceptable levels of timeliness between the data sources and mobile users. In order to provide QoS guarantees in different application scenarios and contexts, it is necessary to model the system performance by incorporating the intermittent connectivity. Queueing Network Models (QNMs) offer a simple modeling environment, which can be used to represent various application scenarios, and provide accurate analytical solutions for performance metrics, such as system response time. We provide an analytical solution regarding the end-to-end response time between the users and the

data sources by modeling the intermittent connectivity of mobile users with product-form QNMs. We utilize the publish/subscribe middleware as the underlying communication infrastructure for the mobile users. To represent the subscriber's connections/disconnections, we model and solve analytically an ON/OFF queueing system by applying a mean value approach. Finally, we validate our model using both simulations with real-world workload traces and comparison with an actual implementation of a Java Messaging Service middleware. The deviations between the performance results foreseen by the analytical model and the ones provided by the simulator and the prototype implementation of a real system are shown to be less than 5% for a variety of scenarios.

- Large-scale mobile environments are characterized by, among others, a large number of mobile users, intermittent connectivity and non-homogeneous arrival rate of data to the users, depending on the region's context. Multiple application scenarios in major cities need to address the above situation for the creation of robust mobile systems. Towards this, it is fundamental to enable system designers to tune a communication infrastructure using various parameters depending on the specific context. We take a first step towards enabling an application platform for large-scale information management relying on mobile social crowd-sourcing [26]. To inform the stakeholders of expected loads and costs, we model a large-scale mobile pub/sub system as a queueing network. We introduce additional timing constraints such as (i) mobile user's intermittent connectivity period; and (ii) data validity lifetime period (e.g. that of sensor data). Using our MobileJINQS simulator (<http://xsb.inria.fr/d4d#mobilejinqs>), we parameterize our model with realistic input loads derived from the D4D CDR (Call Detail Record) dataset (<http://www.d4d.orange.com/en/home>) and varied lifetime periods in order to analyze the effect on response time. This work provides system designers with coarse grain design time information when setting realistic loads and time constraints [18].

7.3. Urban Civics: An IoT Middleware for Democratizing Crowdsensed Data in Smart Societies

Participants: Valérie Issarny, Fadwa Rebhi, Animesh Pathak, Sara Hachem.

The growth of our cities comes along with the aggravation of urban nuisances (e.g., air pollution), which significantly alters the citizens' quality of life and especially their health. It then becomes essential to ensure the growth of cities is both environmentally and socially sustainable. As computer scientists, it is our vision that ICT shall play a key role in achieving the above sustainability requirements, as already put forward by the smart city/society concept. However, smart cities have mostly emphasized the big data dimension and related knowledge engineering to ease the management of the city's infrastructure and resources. While this is an important part of smart cities, we believe that ICT should be leveraged to promote participatory democracy so that citizens and government can communicate openly about the issues facing their societies as much as about their solutions. Toward that goal, we have introduced the Urban Civics middleware, which addresses three complementary research questions underlying participatory democracy from an ICT perspective [20], [21]:

(RQ1) How to leverage the richness of urban sensors of the new digital era that features the Internet of Things, open data, social networking, and mobile computing to serve both citizens and government with better insights? Our answer lies in connecting those various data sources where probabilistic protocols combined with semantic technology allow for an urban-scale middleware solution.

(RQ2) How to assimilate urban data so as to generate explanatory city models to inform urban problem solving? Our solution leverages data assimilation (developed by the Inria CLIME team) that has proven successful in geosciences and paves the way to the comprehensive integration of heterogeneous data sources whose accuracy may vary significantly.

(RQ3) How to integrate the solutions to the above into a scalable urban middleware and further ensure citizen participation? Building on our past experience in developing middleware solutions for the mobile environment and especially the – mobile – Internet of Things, we have conceived and introduced the architecture of Urban Civics, a novel IoT middleware solution for democratizing

crowd-sensed data in smart societies. We are in particular confident that, in addition to leveraging existing incentive mechanisms, the citizen participation will also be prompted by the very nature of participatory democracy. However, such an assumption needs to be validated through actual experiments at an urban scale for which we deploy use cases in the Paris and San Francisco Bay areas.

7.4. AppCivist: Engineering Software Assemblies for Participatory Democracy

Participants: Valérie Issarny, Cristhian Parra Trepowski, Animesh Pathak.

Information and communication technologies (ICT) are profoundly changing the nature of human social and environmental interactions. One such change concerns innovations in the way that citizens both interact with government institutions and engage in greater self-government through democratic assembly and collective action. Our research focuses on this transformation of politics, asking how new social media can contribute to new forms of democracy. The pervasive use of ICT suggests that they present an unprecedented opportunity to rethink the constraints of time and space that are generally thought to make the exercise of a more direct and engaging democracy at a large scale practically impossible. In effect, ICT challenge the assumption that citizens of large political units must be content with systems of representative democracy that typically produce a more passive and legalistic citizenship than an active and participatory one.

To consider this challenge, we undertake a pragmatic and modest investigation of how ICT and more precisely software systems can contribute to enabling direct democracy at a large scale. Our research has two immediate objectives. One is to engineer software that leverages the reach of the Internet and the powers of computation to enhance the experience and efficacy of civic participation. The second is to use the ICT software platform to induce the associational forms of a new digitally-inspired citizenship among residents.

Our research is multi-disciplinary in nature, bringing together anthropologists and computer scientists to coinvestigate how to build software systems that promote the development of such digital democratic assemblies and citizens. Our initiative is further rooted in the principles of social activism in that we want to provide citizens with new software systems that help them articulate projects, deliberate directly among themselves, and mobilize activities. A number of digital tools and in particular social networks and web-based content management systems already support aspects of social activism. However, these tools need to be customized as much as composed to become really useful for activists. To that end, we have set the principles of the AppCivist service-oriented software platform in [24]. AppCivist is built around the vision of letting activist users compose their own applications, called Assemblies, using relevant Internet-based components that enable various aspects of democratic assembly and collective action. Starting from a social science perspective, we identified the following high-level categories of functions for AppCivist Assemblies: Mobilizing people, Co-creating proposals, Acting collectively, and Communicating.

Following, we have concentrated on developing the first instance of AppCivist for Participatory Budgeting (PB), as a representative use case of participatory democracy. As a result, we are able to account for various initiatives in citizen participation, including lessons learned from existing PB campaigns worldwide since their emergence in Brazil in the late 1980s. Research contributions more specifically relate to [22]:

- *State of the art survey and analysis of software systems that contribute to enabling participatory democracy*, which lacks an adequate bottom-up approach to digital proposal making. Such an approach would allow groups of citizens to self-assemble on the basis of common interests and enable the resulting citizen assemblies to initiate ideas and elaborate on them using convenient assemblies of software services.
- *State of the art survey and analysis of digital tools oriented towards Participatory Budgeting*, where leveraging ICT to enable truly urban-scale participation in PB campaigns remains unrealized. AppCivist-PB utilizes the concepts of *citizen assembly* and *software assembly* to address this challenge.
- *AppCivist-PB software architecture* enabling citizen and software assemblies, which following the design of AppCivist introduced in [24] strictly adheres to the principles of service orientation. In

that framework, citizen assemblies allow registered users and groups of users to self assemble into higher-level groups to coordinate idea generation and to elaborate proposals through versioning. In a complementary way, software assemblies adhere to the well-known principle of service composition, configuring software services and components oriented towards the implementation of functions supporting participatory democracy.

- *AppCivist-PB prototype* permits an early assessment of the effectiveness of AppCivist-PB in supporting actual urban-scale PB campaigns, such as the one of Paris in 2015. In addition, the prototype provides an opportunity to experiment with developing service wrappers to integrate third-party services (e.g., Etherpad.org) into its software assemblies. In the near future, we intend to automate this integration as much as possible, building on our background in the synthesis of mediators [13], [12].

This research is carried out in collaboration with the Social Apps Lab at CITRIS at UC Berkeley in the context of CityLab@Inria and Inria@SiliconValley.

MOAIS Project-Team

5. New Results

5.1. TABARNAC: Tools for Analyzing Behavior of Applications Running on NUMA Architecture.

In modern parallel architectures, memory accesses represent a common bottleneck. We develop TABARNAC, a tool for analyzing the memory behavior of parallel applications with a focus on NUMA architectures. TABARNAC provides a new visualization of the memory access behavior, focusing on the distribution of accesses by thread and by structure. Such visualization allows the developer to easily understand why performance issues occur. Using TABARNAC, we explain why some applications do not benefit from data and thread mapping. Moreover, we propose several code modifications to improve the memory access behavior of several parallel applications [29].

5.2. Computing the Rank Profile Matrix

We propose the definition of a new matrix invariant, the rank profile matrix, summarizing all information on the row and column rank profiles of all the leading sub-matrices. We also explore the conditions for a Gaussian elimination algorithm to compute all or part of this invariant, through the corresponding PLUQ decomposition [12].

5.3. Parallel Algebraic Linear Algebra Dedicated Interface

We propose a domain specific language based on C/C++ macros, PALADIn (Parallel Algebraic Linear Algebra Dedicated Interface) [15]. This domain specific language allows the user to write C++ code and benefits from sequential and parallel executions on shared memory architectures. With a unique syntax, the user can switch between different parallel runtime systems such as OpenMP, TBB and xKaapi. This interface provides data and task parallelism and has been used for recursion-based parallelization of exact dense linear algebra routines[7].

5.4. Communication models insights meet simulations

It is well-known that taking into account communications while scheduling jobs in large scale parallel computing platforms is a crucial issue. In modern hierarchical platforms, communication times are highly different when occurring inside a cluster or between clusters. Thus, allocating the jobs taking into account locality constraints is a key factor for reaching good performances. However, several theoretical results prove that imposing such constraints reduces the solution space and thus, possibly degrades the performances. In practice, such constraints simplify implementations and most often lead to better results. Our aim in this work is to bridge theoretical and practical intuitions, and check the differences between constrained and unconstrained schedules (namely with respect to locality and node contiguity) through simulations. We have developed a generic tool, using SimGrid as the base simulator, enabling interactions with external batch schedulers to evaluate their scheduling policies. The results confirm that insights gained through theoretical models are ill-suited to current architectures and should be reevaluated [13].

5.5. Adaptive Resource and Job Management for Limited Power Consumption

The last decades have been characterized by an evergrowing requirement in terms of computing and storage resources. This tendency has recently put the pressure on the ability to efficiently manage the power required to operate the huge amount of electrical components associated with state-of-the-art high performance computing systems. The power consumption of a supercomputer needs to be adjusted based on varying power budget or electricity availabilities. As a consequence, Resource and Job Management Systems have to be adequately adapted in order to efficiently schedule jobs with optimized performance while limiting power usage whenever needed. We introduce in this paper a new scheduling strategy that can adapt the executed workload to a limited power budget. The originality of this approach relies upon a combination of speed scaling and node shutdown techniques for power reductions. It is implemented into the widely used resource and job management system SLURM. Finally, it is validated through large scale emulations using real production workload traces of the supercomputer Curie [17].

5.6. Lessons Learned from Building In Situ Coupling Frameworks

Over the past few years, the increasing amounts of data produced by large-scale simulations have motivated a shift from traditional offline data analysis to in situ analysis and visualization. In situ processing began as the coupling of a parallel simulation with an analysis or visualization library, motivated primarily by avoiding the high cost of accessing storage. Going beyond this simple pairwise tight coupling, complex analysis workflows today are graphs with one or more data sources and several interconnected analysis components. In this paper, we review four tools that we have developed to address the challenges of coupling simulations with visualization packages or analysis workflows: Damaris, Decaf, FlowVR and Swift. This self-critical inquiry aims to shed light not only on their potential, but most importantly on the forthcoming software challenges that these and other in situ analysis and visualization frameworks will face in order to move toward exascale [11]. Besides, focusing on asynchronous In Situ Processing with Gromacs, we have exhibited how to take Advantage of GPUs [25].

5.7. Design and analysis of scheduling strategies for multi-CPU and multi-GPU architectures

In [8], we present a comparison of scheduling strategies for heterogeneous multi-CPU and multi-GPU architectures. We designed and evaluated four scheduling strategies on top of XKaapi runtime: work stealing, data-aware work stealing, locality-aware work stealing, and Heterogeneous Earliest-Finish-Time (HEFT). On a heterogeneous architecture with 12 CPUs and 8 GPUs, we analysed our scheduling strategies with four benchmarks: a BLAS-1 AXPY vector operation, a Jacobi 2D iterative computation, and two linear algebra algorithms Cholesky and LU. We conclude that the use of work stealing may be efficient if task annotations are given along with a data locality strategy. Furthermore, our experimental results suggests that HEFT scheduling performs better on applications with very regular computations and low data locality.

MUSE Team

6. New Results

6.1. Home Network or Access Link? Locating Last-mile Downstream Throughput Bottlenecks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

As home networks see increasingly faster downstream throughput speeds, a natural question is whether users are benefiting from these faster speeds or simply facing performance bottlenecks in their own home networks. In our paper recently accepted for publication in PAM'16, we studied whether downstream throughput bottlenecks occur more frequently in their home networks or in their access ISPs. We identified lightweight metrics that can accurately identify whether a throughput bottleneck lies inside or outside a user's home network and developed a detection algorithm that locates these bottlenecks. We validated this algorithm in controlled settings and characterized bottlenecks on two deployments, one of which included 2,652 homes across the United States. We found that wireless bottlenecks are more common than access-link bottlenecks—particularly for home networks with downstream throughput greater than 20 Mbps, where access-link bottlenecks are relatively rare.

6.2. On the Reliability of Profile Matching Across Large Online Social Networks

Participants: Oana Goga and Krishna Gummadi (MPI-SWS), Patrick Loiseau (EURECOM), Robin Sommer (ICSI), Renata Teixeira

Matching the profiles of a user across multiple online social networks brings opportunities for new services and applications as well as new insights on user online behavior, yet it raises serious privacy concerns. Prior literature has showed that it is possible to accurately match profiles, but their evaluation focused only on sampled datasets. In our KDD'15 paper [2], we study the extent to which we can reliably match profiles in practice, across real-world social networks, by exploiting public attributes, i.e., information users publicly provide about themselves. Today's social networks have hundreds of millions of users, which brings completely new challenges as a reliable matching scheme must identify the correct matching profile out of the millions of possible profiles. We first define a set of properties for profile attributes—Availability, Consistency, non-Impersonability, and Discriminability (ACID)—that are both necessary and sufficient to determine the reliability of a matching scheme. Using these properties, we propose a method to evaluate the accuracy of matching schemes in real practical cases. Our results show that the accuracy in practice is significantly lower than the one reported in prior literature. When considering entire social networks, there is a non-negligible number of profiles that belong to different users but have similar attributes, which leads to many false matches. Our paper sheds light on the limits of matching profiles in the real world and illustrates the correct methodology to evaluate matching schemes in realistic scenarios.

6.3. Exploiting crowd sourced reviews to explain movie recommendation

Participants: Sara El Aouad, Christophe Dupuy, Francis Bach, and Renata Teixeira (Inria), Christophe Diot (Technicolor)

Streaming services such as Netflix, M-Go, and Hulu use advanced recommender systems to help their customers identify relevant content quickly and easily. These recommenders display the list of recommended movies organized in sublists labeled with the genre or some more specific labels. Unfortunately, existing methods to extract these labeled sublists require human annotators to manually label movies, which is time-consuming and biased by the views of annotators. In our work [6], we design a method that relies on crowd-sourced reviews to automatically identify groups of similar movies and label these groups. Our method takes the content of movie reviews available online as input for an algorithm based on Latent Dirichlet Allocation (LDA) that identifies groups of similar movies. We separate the set of similar movies that share the same combination of genre in sublists and personalize the movies to show in each sublist using matrix factorization. The results of a side-by-side comparison of our method against Technicolor's M-Go VoD service are encouraging.

6.4. Characterizing Home Device Usage From Wireless Traffic Time Series

Participants: Katsiaryna Mirylenka (IBM Research - Zurich), Vassilis Christophides, Themis Palpanas (Paris Descartes University), Ioannis Pefkianakis (Hewlett Packard Labs), Martin May (Technicolor).

The analysis of *temporal behavioral patterns* of home network users can reveal important information to Internet Service Providers (ISPs) and help them to optimize their networks and offer new services (e.g., remote software upgrades, troubleshooting, energy savings). Our study [4] uses time series analysis of continuous traffic data from wireless home networks, to extract traffic patterns recurring within, or across homes, and *assess the impact of different device types (fixed or portable) on home traffic*. Traditional techniques for time series analysis are not suited in this respect, due to the limited stationary and evolving distribution properties of wireless home traffic data. We propose a novel framework that relies on a *correlation-based similarity* measure of time series, as well as a notion of *strong stationarity* to define recurring motifs and dominant devices. Using this framework, we analyze the wireless traffic collected from 196 home gateways over two months. Our framework goes beyond existing application-specific analysis techniques, such as analysis of wireless traffic, which mainly rely on data aggregated across hundreds, or thousands of users. It enables the extraction of recurring patterns from traffic time series of individual homes, leading to a much more fine-grained analysis of the behavior patterns of the users. We also determine the best time aggregation policy w.r.t. to the number and statistical importance of the extracted motifs, as well as the device types dominating these motifs and the overall gateway traffic. Our results show that ISPs can exceed the simple observation of the aggregated gateway traffic and better understand their networks.

6.5. On Continuous Top-k Queries with Real-Time Scoring Functions

Participants: Nelly Vouzoukidou (Google, France), Bernd Amann (LIP6), Vassilis Christophides.

Modern news sharing and social media platforms allow millions of users to *produce and consume information in real-time*. To assess relevance of published information in this new setting, batch scoring based on content similarity, link centrality or page views is no longer sufficient. Instead, streams of events like “replies” (for posting comments), “likes” (for rating content) or “retweets” (for diffusing information) explicitly provided by users represent valuable online feedback on published information that has to be exploited in order to adjust in real-time any available score of information items. Note that in the future Internet of Things (IoT), not only digital, but also physical objects will be expected to be ranked in a fully automated way with respect to real-time human activities (viewing concentration), vital signals (emotional arousal), etc.

Rather than indexing as quickly as possible information items to re-evaluate *snapshot queries*, publish/subscribe systems index *continuous queries* and update on the fly their results each time a new matching item arrives. Existing publish/subscribe systems rely on two alternative continuous filtering semantics, namely *predicate-based* filtering or *similarity-based top-k* filtering. In predicate-based systems, incoming items that match the filtering predicates are simply added to the result list of continuous queries, while in similarity-based top-k publish/subscribe systems, matching items have also to exhibit better relevance w.r.t. the items already appearing as the top-k results of the continuous query. In top-k publish/subscribe systems the relevance of an item remains constant during a pre-specified time window, and once its lifetime exceeds the

item simply expires. Only recently, information recency has become part of the relevance score of continuous queries. Clearly, when information relevance decays as time passes both (a) results lists maintenance and (b) early pruning of the query index traversal are challenged. While these problems have been studied for (textual or spatio-textual) content scoring functions with time decay, non-homogeneous scoring functions accommodating various forms of *query-dependent* and *query-independent* information relevance with time decay is supported only by MeowsReader. In this work we are going beyond this general form of *time-decayed static scores* and consider continuous queries featuring *real-time scoring functions* under the form of *time decaying positive user feedback* for millions of online social media events per minute and millions of user queries.

MYRIADS Project-Team

7. New Results

7.1. Cloud Resource Management

Participants: Ancuta Iordache, Christine Morin, Ghada Moualla, Guillaume Pierre, Matthieu Simonin, Lodewijck Vogelzang.

7.1.1. Application Performance Modeling in Heterogeneous Cloud Environments

Participants: Ancuta Iordache, Lodewijck Vogelzang, Guillaume Pierre.

Heterogeneous cloud platforms offer many possibilities for applications to make fine-grained choices over the types of resources they execute on. This opens for example opportunities for fine-grained control of the tradeoff between expensive resources likely to deliver high levels of performance, and slower resources likely to cost less. We designed a methodology for automatically exploring this performance vs. cost tradeoff when an arbitrary application is submitted to the platform. Thereafter, the system can automatically select the set of resources which is likely to implement the tradeoff specified by the user. We significantly improved the speed at which the system can characterize the performance of an arbitrary application. A first publication on this topic has been published [26], and a second one is in preparation.

7.1.2. Heterogeneous Resource Management

Participants: Ancuta Iordache, Guillaume Pierre.

During her internship at Maxeler Technologies, Ancuta Iordache developed an original technique for virtualizing FPGAs such that they can be used as high-performance computing devices in cloud infrastructures. Virtual FPGAs can be accessed remotely by any VM in the system. They can span multiple physical FPGA, they are elastic, and they can also be shared between multiple tenants. A publication on this topic is currently under evaluation.

7.1.3. Self-adaptatable Hadoop Virtual Clusters

Participants: Christine Morin, Ghada Moualla, Matthieu Simonin.

In the context of Ghada Moualla's Master internship, we designed the Elastic MapReduce Adaptation (EMRA) system to execute Hadoop MapReduce applications with user-defined deadlines in cloud virtual clusters. EMRA integrates an algorithm to automatically adapt the Hadoop cluster size at runtime in order to meet user-defined deadlines. We proposed an automatic scaling algorithm, which monitors the progress of the Map phase of the application during its execution and estimate if the user-defined deadline can be met. If the current allocated resources are not sufficient to meet the deadline, more resources are provisioned. The adaptation service comprises of three main components: (i) a monitor to check the progress of the running application, (ii) an estimator to predict the time needed to complete the application based on its current progress ; (iii) a controller to adapt the size of the virtual cluster by adding virtual machines as needed. The controller takes into account the start-up overhead of the new virtual machines and the time needed for the VM to fetch their input data from the original nodes over the network in order to start their map tasks. We implemented a prototype of the EMRA system in the context of Sahara, an environment for managing Hadoop virtual clusters on top of OpenStack IaaS clouds. We experimented the EMRA system on Grid'5000 with traditional MapReduce benchmarks. We evaluated the relative error of the estimator, the cost for scaling up or down a virtual cluster and showed that the proposed adaptation algorithm allows user-defined deadlines to be met.

7.2. Distributed Cloud Computing

Participants: Teodor Crivat, Yvon Jégou, Vlad Mirel, Christine Morin, Anne-Cécile Orgerie, Edouard Outin, Nikolaos Parlavantzas, Jean-Louis Pizat, Guillaume Pierre, Aboozar Rajabi, Carlos Ruiz Diaz, Arnab Sinha, Genc Tato, Cédric Tedeschi.

7.2.1. A multi-objective adaptation system for the management of a Distributed Cloud

Participants: Yvon Jégou, Edouard Outin, Jean-Louis Pazat.

In this project, we consider a “Distributed Cloud” made of multiple data/computing centers interconnected by a high speed network and belonging to the same administration domain. Moreover, in the Cloud organization targeted here, the network capabilities can be dynamically configured in order to guarantee QoS for streaming or to negotiate bandwidth for example.

As a first step, we are focusing on a single centralised Cloud.

Due to the dynamic capabilities of the Clouds, often referred to as elasticity, there is a strong need to dynamically adapt both platforms and applications to users needs and environmental constraints such as electrical power consumption.

We address the management of a Cloud in order to consider both optimization for energy consumption and for users’ QoS needs. The objectives of this optimization will be negotiated as contracts on Service Level Agreement (SLA). A special emphasis will be put on the distributed aspect of the platform and include both servers and network adaptation capabilities.

The design of the system relies on self-* techniques and on adaptation mechanisms at any level (from IaaS to SaaS). The MAPE-k framework (Monitor-Analysis-Planning-Execution based on knowledge) is used for the implementation of the system. The technical developments are based on the Openstack framework.

We have implemented a system that uses a genetic algorithm to optimize Cloud energy consumption and machine learning techniques to improve the fitness function regarding a real distributed cluster of servers. We have carried out experiments on the OpenStack platform to validate our solution. This experimentation shows that the machine learning produces an accurate energy model, predicting precise values for the simulation.

We are currently refining this model and comparing it to real measurements on the platform.

This work is done in cooperation with the DIVERSE team and in cooperation with Orange under the umbrella of the B-COM Technology Research Center.

7.2.2. Dynamic reconfiguration for multi-cloud applications

Participants: Nikolaos Parlavantzas, Aboozar Rajabi, Carlos Ruiz Diaz, Arnab Sinha.

In the context of the PaaSage European project, we are working on model-based, continuous self-optimization of multi-cloud applications. In particular, we are developing a dynamic adaptation system, capable of transforming the currently running application configuration into a target configuration in a cost-effective and safe manner. In 2015, we have improved and extended the Adapter prototype [45]. The system now fully supports dynamic configuration, including detecting changes, generating reconfiguration plans, validating plans based on a cost-benefit calculation, and executing plans in parallel, improving adaptation performance. Moreover, we have performed initial investigations on the use of PaaSage for supporting Internet of Things (IoT) applications [27]. Finally, in the context of Carlos Ruiz’s stay, we are defining a model for managing the configuration of cloud applications and environments. This model is based on feature modeling and the derived configurations are mapped to PaaSage models.

7.2.3. Towards a distributed cloud inside the backbone

Participants: Christine Morin, Anne-Cécile Orgerie, Genc Tato, Cédric Tedeschi.

The DISCOVERY proposal officially started at the end of 2015. It is an Inria Project Lab (IPL) led by Adrien Lebre from the ASCOLA team, and currently on leave at Inria. It aims at designing a distributed cloud, leveraging the resources we can find in the network backbone.⁰ In practice, this work is intended to get integrated within the OpenStack software <https://www.openstack.org/> so as to decentralize its whole architecture.

⁰The DISCOVERY website: <http://beyondtheclouds.github.io>

In this context, and in collaboration with ASCOLA and ASAP teams, we started the design of an overlay network whose purpose is to be able, with a limited cost, to locate geographically-close nodes from any point of the network. In this framework, the PhD thesis of Genc Tato started in December 2015. It aims at developing locality mechanisms at the data management layer.

We have also started an energy/cost-benefit analysis of a decentralized Cloud infrastructure like the one proposed within Discovery. This work is conducted by Anthony Simonet, a post-doctoral researcher on an Inria contract for the Discovery IPL and co-supervised by Adrien Lebre from the ASCOLA team and Anne-Cécile Orgerie from Myriads team.

7.2.4. Mobile edge cloud computing with ConPaaS

Participants: Teodor Crivat, Vlad Mirel, Guillaume Pierre.

Interactive multi-user applications usually rely on intermediate cloud servers to mediate the inter-user interaction. However, current mobile networks exhibit network latencies in the order of 50-150 ms between the device and any cloud. Such latencies make it impossible to create smooth interactions with the end user. To enable an “instantaneous” feeling, augmented reality applications require that end-to-end latencies should remain below 20 ms.

To address these issues, we extended ConPaaS to support the deployment of cloud applications in a distributed set of Raspberry Pi machines. The motivation is to reduce the latency compared to a traditional deployment where the backend is located in an external cloud: instead of reaching the cloud through a wide-area network, in this setup each cloud node is also equipped with a wifi hotspot which allows local users to access it directly.

7.2.5. Fog Computing

Participant: Jean-Louis Pazat.

The concept of “Fog Computing” is currently developed on the idea of hosting instances of services not on centralized datacenters (i.e. the “Cloud”), but on a highly distributed infrastructure: the Internet Edge (i.e. the “Fog”). This infrastructure consists in geographically distributed computing resources with relatively small capabilities. Compared with datacenters, a “Fog” infrastructure is able to offer to Service Providers a shorter distance from the service to the user but with the same flexibility of software deployment and management.

This work focus on the problem of resource allocation in such infrastructure when considering services in the area of Internet of Things, Social Networks or Online Gaming. For such use-cases, service-to-user latency is a critical parameter for the quality of experience. Optimizing such a parameter is an objective for the platform built on top of the Fog Infrastructure that will be dedicated to the deployment of the considered service. In order to achieve such a goal, the platform needs to select some strategies for the allocation of network and computing resources, based on the initial requirements for service distribution.

We are designing a prototype based on micro services and we are considering low overhead virtualization systems using containers. This prototype is intended to run inside an Internet Box or inside a LAN disk server at user’s home. The whole system will be intended to be used very small or medium size user communities willing to share devices and data. The main characteristics of the system will be reliable distributed storage and distributed execution of services.

This work is part of Bruno Stevant’s PhD thesis, which began in December 2014. It is done in cooperation with the REOP team, Institut Mines telecom/IRISA.

7.3. Cloud Security

Participants: Anna Giannakou, Christine Morin, Jean-Louis Pazat, Louis Rilling, Amir Teshome Wonjiga.

7.3.1. Security Monitoring of Clouds

Participants: Anna Giannakou, Christine Morin, Jean-Louis Pazat, Louis Rilling, Amir Teshome Wonjiga.

We aim at making security monitoring a dependable service for IaaS cloud customers. To this end, we study three topics:

- defining relevant SLA terms for security monitoring,
- enforcing and evaluating SLA terms,
- making the SLA terms enforcement mechanisms self-adaptable to cope with the dynamic nature of clouds.

The considered enforcement and evaluation mechanisms should have a minimal impact on performance.

In 2015 we started to study the state of the art about SLA for security monitoring in clouds, as well as about evaluating security monitoring setups in clouds.

In 2015 we also studied the self-adaptation issues of security monitoring with two kinds of security monitoring components: a network intrusion detection system (NIDS), and a secured application-level firewall. Moreover a new approach to secure an application-level firewall has been proposed.

To experiment with both kinds of components, a prototype called SAIDS has been implemented in the OpenStack-based IaaS cloud testbed that was setup in 2014. The NIDS software used is Snort. The application-level firewall is based on Linux nftables and Open vSwitch. In order to study more complex security monitoring setups, SAIDS will be extended in 2016.

A preliminary evaluation of SAIDS has been published in the doctoral symposium of CCGrid 2015. A more complete evaluation of SAIDS as well as the evaluation of the application-level firewall will be done in 2016.

7.4. Greening Clouds

Participants: Maria Del Mar Callau Zori, Ismael Cuadrado Cordero, David Guyon, Sabbir Hasan Rochi, Yunbo Li, Christine Morin, Anne-Cécile Orgerie, Jean-Louis Pazat, Guillaume Pierre.

7.4.1. Energy-aware IaaS-PaaS co-design

Participants: Maria Del Mar Callau Zori, Anne-Cécile Orgerie, Guillaume Pierre.

The wide adoption of the cloud computing paradigm plays a crucial role in the ever-increasing demand for energy-efficient data centers. Driven by this requirement, cloud providers resort to a variety of techniques to improve energy usage at each level of the cloud computing stack. However, prior studies mostly consider resource-level energy optimizations in IaaS clouds, overlooking the workload-related information locked at higher levels, such as PaaS clouds. We conducted an extensive experimental evaluation of the effect of a range of Cloud infrastructure operations (start, stop, migrate VMs) on their computing throughput and energy consumption, and derived a model to help drive cloud reconfiguration operations according to performance/energy requirements. A publication on this topic is in preparation.

7.4.2. Energy-efficient cloud elasticity for data-driven applications

Participants: David Guyon, Anne-Cécile Orgerie, Christine Morin.

Distributed and parallel systems offer to users tremendous computing capacities. They rely on distributed computing resources linked by networks. They require algorithms and protocols to manage these resources in a transparent way for users. Recently, the maturity of virtualization techniques has allowed for the emergence of virtualized infrastructures (Clouds). These infrastructures provide resources to users dynamically, and adapted to their needs. By benefiting from economies of scale, Clouds can efficiently manage and offer virtually unlimited numbers of resources, reducing the costs for users.

However, the rapid growth for Cloud demands leads to a preoccupying and uncontrolled increase of their electric consumption. In this context, we will focus on data driven applications which require to process large amounts of data. These applications have elastic needs in terms of computing resources as their workload varies over time. While reducing energy consumption and improving performance are orthogonal goals, this internship aims at studying possible trade-offs for energy-efficient data processing without performance impact. As elasticity comes at a cost of reconfigurations, these trade-offs will consider the time and energy required by the infrastructure to dynamically adapt the resources to the application needs.

The master internship work of David Guyon on this topic has been presented at IEEE GreenCom 2015 [39]. This work will be continued during David's PhD thesis.

7.4.3. *Energy-efficient and network-aware resource allocation in Cloud infrastructures*

Participants: Ismael Cuadrado Cordero, Christine Morin, Anne-Cécile Orgerie.

Energy consumption in cloud computing has become a key environmental and economic concern. Our work aims at designing energy-efficient resource allocation for Cloud infrastructures. The ever-growing appetite of new applications for network resources leads to an unprecedented electricity bill, and for these bandwidth-hungry applications, networks can become a significant bottleneck. New algorithms have to be designed integrating the data locality dimension to optimize computing resource allocation while taking into account the fluctuating limits of network resources. Towards this end, we proposed GRaNADA, a semi-decentralized Platform-as-a-service (PaaS) architecture for real-time multiple-users applications. Our architecture geographically distributes the computation among the clients of the cloud, moving the computation away from the datacenter to save energy - by shutting down or downgrading non utilized resources such as routers and switches, servers, etc. - and provides lower latencies for users. GRaNADA implements the concept of micro-cloud, a fully autonomous energy-efficient subnetwork of clients of the same service, designed to keep the greenest path between its nodes. Along with GRaNADA, we proposed DEEPACC, a cloud-aware routing protocol which distributes the connection between the nodes. Our system GRaNADA targets services where the geographical distribution of clients working on the same data is limited - for example, a shared on-line document - or those services where, even if the geographical distribution of clients is high, the upload data communication to the cloud is small - for instance a light social network like Twitter. We compared our approach with two main existing solutions - replication of data in the edge and traditional centralized cloud computing. Our approach based on micro-clouds exhibits interesting properties in terms of QoS and especially latency. Simulations show that, using the proposed PaaS, one can save up to 75% of the spent network energy compared to traditional centralized cloud computing approaches. Our approach is also more energy-efficient than the most popular semi-decentralized solutions, like nano data centers. This work has been presented at IEEE GreenCom 2015 [18].

We also evaluated the suitability of using micro-clouds in the context of smart cities. We investigated the idea to build a local cloud on top of networking resources spread across a defined area and including the mobile devices of the users. This local cloud is managed by lightweight mechanisms in order to handle users who can appear/disappear and move. We used a scenario considering a platform for neighborhood services and showed that micro-clouds make better use of the network, reducing the amount of unnecessary data traveling through external networks. This work is currently under review for a conference.

7.4.4. *Resource allocation in a Cloud partially powered by renewable energy sources*

Participants: Yunbo Li, Anne-Cécile Orgerie.

We propose here to design a disruptive approach to Cloud resource management which takes advantage of renewable energy availability to perform opportunistic tasks. To begin with, the considered Cloud is mono-site (i.e. all resources are in the same physical location) and performs tasks (like web hosting or MapReduce tasks) running in virtual machines. This Cloud receives a fixed amount of power from the regular electric Grid. This power allows it to run usual tasks. In addition, this Cloud is also connected to renewable energy sources (such as windmills or solar cells) and when these sources produce electricity, the Cloud can use it to run more tasks.

The proposed resource management system needs to integrate a prediction model to be able to forecast these extra-power periods of time in order to schedule more work during these periods. Batteries will be used to guarantee that enough energy is available when switching on a new server working exclusively on renewable energy. Given a reliable prediction model, it is possible to design a scheduling algorithm that aims at optimizing resource utilization and energy usage, problem known to be NP-hard. The proposed heuristics will thus schedule tasks spatially (on the appropriate servers) and temporally (over time, with tasks that can be planned in the future).

This work is done in collaboration with Ascola team from LINA in Nantes. Two publications have been accepted this year on this topic for: SmartGreens 2015 [15] and IEEE GreenCom 2015 [21].

7.4.5. *SLA driven Cloud Auto-scaling for optimizing energy footprint*

Participants: Sabbir Hasan Rochi, Jean-Louis Pazat.

As a direct consequence of the increasing popularity of Internet and Cloud Computing services, data centers are amazingly growing and hence have to urgently face energy consumption issues. At the Infrastructure-as-a-Service (IaaS) layer, Cloud Computing allows to dynamically adjust the provision of physical resources according to Platform-as-a-Service (PaaS) needs while optimizing energy efficiency of the data center.

The management of elastic resources in Clouds according to fluctuating workloads in the Software-as-a-Service (SaaS) applications and different Quality-of-Service (QoS) end-user's expectations is a complex issue and cannot be done dynamically by a human intervention. We advocate the adoption of Autonomic Computing (AC) at each XaaS layer for responsiveness and autonomy in front of environment changes. At the SaaS layer, AC enables applications to react to a highly variable workload by dynamically adjusting the amount of resources in order to keep the QoS for the end users. Similarly, at the IaaS layer, AC enables the infrastructure to react to context changes by optimizing the allocation of resources and thereby reduce the costs related to energy consumption. However, problems may occur since those self-managed systems are related in some way (e.g. applications depend on services provided by a cloud infrastructure): decisions taken in isolation at given layer may interfere with other layers, leading whole system to undesired states.

We have defined a scheme for green energy management in the presence of explicit and implicit integration of renewable energy in datacenter [13]. More specifically we propose three contributions: i) we introduce the concept of virtualization of green energy to address the uncertainty of green energy availability, ii) we extend the Cloud Service Level Agreement (CSLA) language to support Green SLA introducing two new threshold parameters and iii) we introduce greenSLA algorithm which leverages the concept of virtualization of green energy to provide per interval specific Green SLA. Experiments were conducted with real workload profile from PlanetLab and server power model from SPECpower to demonstrate that, Green SLA can be successfully established and satisfied without incurring higher cost.

This work is done in collaboration with Ascola team from LINA in Nantes.

7.5. Energy-efficient Computing Infrastructures

Participants: Christine Morin, Anne-Cécile Orgerie, Martin Quinson.

7.5.1. *Simulating the impact of DVFS within SimGrid*

Participants: Christine Morin, Anne-Cécile Orgerie, Martin Quinson.

Simulation is a popular approach for studying the performance of HPC applications in a variety of scenarios. However, simulators do not typically provide insights on the energy consumption of the simulated platforms. Furthermore, studying the impact of application configuration choices on energy is a difficult task, as not many platforms are equipped with the proper power measurement tools. The goal of this work is to enable energy-aware experimentation within the SimGrid simulation toolkit, by introducing a model of application energy consumption and enabling the use of Dynamic Voltage and Frequency Scaling (DVFS) techniques for the simulated platforms. We provide the methodology used to obtain accurate energy estimations, highlighting the simulator calibration phase. The proposed energy model is validated by means of a large set of experiments featuring several benchmarks and scientific applications. This work is available in the latest SimGrid release. This work is done in collaboration with the Mescal team from LIG in Grenoble. A paper is currently under preparation on this work.

7.5.2. *Simulating Energy Consumption of Wired Networks*

Participants: Timothée Haudebourg, Anne-Cécile Orgerie.

Predicting the performance of applications, in terms of completion time and resource usage for instance, is critical to appropriately dimensioning resources that will be allocated to these applications. Current applications, such as web servers and Cloud services, require lots of computing and networking resources. Yet, these resource demands are highly fluctuating over time. Thus, adequately and dynamically dimensioning these resources is challenging and crucial to guarantee performance and cost-effectiveness. In the same manner, estimating the energy consumption of applications deployed over heterogeneous cloud resources is important in order to provision power resources and make use of renewable energies. Concerning the consumption of entire infrastructures, some studies show that computing resources represent the biggest part in the Cloud's consumption, while others show that, depending on the studied scenario, the energy cost of the network infrastructure that links the user to the computing resources can be bigger than the energy cost of the servers.

In this work, we aim at simulating the energy consumption of wired networks which receive little attention in the Cloud computing community even though they represent key elements of these distributed architectures. To this end, we are contributing to the well-known open-source simulator ns3 by developing an energy consumption module named ECOFEN.

In 2015, this simulator has been extended to integrate two more green levers: low power idle (IEEE 802.3az) and adaptive link rate. This work has been done during the internship of Timothée Haudebourg (L3 ENS Rennes) and a publication is currently under preparation.

7.5.3. Multicriteria scheduling for large-scale HPC environments

Participant: Anne-Cécile Orgerie.

Energy consumption is one of the main limiting factor for the design and deployment of large scale numerical infrastructures. The road towards "Sustainable Exascale" is a challenge with a target of 50 Gflops per watt. Energy efficiency must be taken into account and must be combined with other criteria like performance, resilience, Quality of Service.

As platforms become more and more heterogeneous (co-processors, GPUs, low power processors...), an efficient scheduling of applications and services at large scale remains a challenge. In this context, we will explore and propose a multicriteria scheduling model and framework for large scale HPC systems. Based on real energy measurements and calibrations, we will propose some performance and energy models and will build a multi criteria scheduler. Simulation on selected scenario will be explored and a prototype will be designed for ensuring experimental validation.

This work is done in collaboration with ROMA and Avalon teams from LIP in Lyon.

7.6. Decentralized and Adaptive workflows

Participants: Jean-Louis Pazat, Javier Rojas Balderrama, Matthieu Simonin, Cédric Tedeschi, Palakiyem Wallah.

7.6.1. Adaptive Workflows with Chemical Computing

Participants: Javier Rojas Balderrama, Matthieu Simonin, Cédric Tedeschi.

We have designed a high-level programming model based on the HOCL rule-based language to express workflow adaptation. It was specifically designed to support changes in the workflow logic at run time. This mechanism was implemented within the GinFlow software and experimented over the Grid'5000 platform. An article was just accepted for publication at the IPDPS 2016 conference.

7.6.2. Best-effort decentralized workflow execution

Participants: Jean-Louis Pazat, Cédric Tedeschi, Palakiyem Wallah.

We are currently proposing a simple workflow model for workflow execution in platforms with limited computing resources and services. The key idea is to devise a best-effort workflow engine that does not require a strong centralized orchestrator. Such a workflow engine relies on point-to-point cooperation between nodes supporting the execution. A minimalistic demonstrator of these concepts has been devised and implemented. Early experiments have been conducted on a single machine.

7.7. Experimental Platforms

Participants: Julien Lefeuvre, David Margery.

7.7.1. Contribution to *Fed4FIRE* testbed

Participants: Julien Lefeuvre, David Margery.

In Fed4FIRE, two key technologies have been adopted as common protocols to enable experimenters to interact with testbeds: Slice Federation Architecture (SFA), to provision resources, and Control and Management Framework for Networking Testbeds (OMF) to control them. In 2015, the main area of work has been the implementation of an SFA API to BonFIRE, still on-going. In the process, we wrote the reference documentation to write a new delegate for geni-tools, the reference implementation of SFA maintained by the GENI project office. This codebase has now been made public on github, in part because of our interactions with the code and suggested changes to ease writing new delegates. We have also contributed to the design of a service layer proxy mechanisms so that testbeds with http based APIs can be queried by any Fed4FIRE user using a standard authentications mechanism. The BonFIRE API has been made available through that mechanism, based on XML documents signed using the XML Signature specification.

PHOENIX Project-Team

7. New Results

7.1. Tablet-Based Activity Schedule in Mainstream Environment for Children with Autism and Children with ID

Including children with Autism Spectrum Disorders (ASD) in mainstreamed environments creates a need for new interventions whose efficacy must be assessed in situ. We present a tablet-based application for activity schedules that has been designed following a participatory design approach involving mainstream teachers, special-education teachers and school aides. This application addresses two domains of activities: classroom routines and verbal communications. We assessed the efficiency of our application with two overlapping user-studies in mainstream inclusion, sharing a group of children with ASD. The first experiment involved 10 children with ASD, where 5 children were equipped with our tablet-based application and 5 were not equipped. We show that (1) the use of the application is rapidly self-initiated (after two months for almost all the participants) and that (2) the tablet-supported routines are better performed after three months of intervention. The second experiment involved 10 children equipped with our application; it shared the data collected for the 5 children with ASD and compared them with data collected for 5 children with Intellectual Disabilities – ID. We show that (1) children with ID are not autonomous in the use of the application at the end of the intervention; (2) both groups exhibited the same benefits on classroom routines; and, (3) children with ID improve significantly less their performance on verbal communication routines. These results are discussed in relation with our design principles. Importantly, the inclusion of a group with another neurodevelopmental condition provided insights about the applicability of these principles beyond the target population of children with ASD.

7.2. Age and active navigation effects on episodic memory: A virtual reality study

We investigated the navigation-related age effects on learning, proactive interference semantic clustering, recognition hits, and false recognitions in a naturalistic situation using a virtual apartment-based task. We also examined the neuropsychological correlates (executive functioning [EF] and episodic memory) of navigation-related age effects on memory. Younger and older adults either actively navigated or passively followed the computer-guided tour of an apartment. The results indicated that active navigation increased recognition hits compared with passive navigation, but it did not influence other memory measures (learning, proactive interference, and semantic clustering) to a similar extent in either age group. Furthermore, active navigation helped to reduce false recognitions in younger adults but increased those made by older adults. This differential effect of active navigation for younger and older adults was accounted for by EF score. Like for the subject-performed task effects, the effects from the navigation manipulation were well accounted for by item-specific/relational processing distinction, and they were also consistent with a source monitoring deficit in older adults.

7.3. Constraining application behaviour by generating languages

Writing a platform for reactive applications which enforces operational constraints is difficult, and has been approached in various ways. In this experience report, we detail an approach using an embedded DSL which can be used to specify the structure and permissions of a program in a given application domain. Once the developer has specified which components an application will consist of, and which permissions each one needs, the specification itself evaluates to a new, tailored, language. The final implementation of the application is then written in this specialised environment where precisely the API calls associated with the permissions which have been granted, are made available. Our prototype platform targets the domain of mobile computing ,

and is implemented using Racket. It demonstrates resource access control (e.g., camera, address book, etc.) and tries to prevent leaking of private data. Racket is shown to be an extremely effective platform for designing new programming languages and their run-time libraries. We demonstrate that this approach allows reuse of an inter-component communication layer, is convenient for the application developer because it provides high-level building blocks to structure the application, and provides increased control to the platform owner, preventing certain classes of errors by the developer.

7.4. A Unifying Notification System To Scale Up Assistive Services

Aging creates needs for assistive technology to support all activities of daily living (meal preparation, dressing, social participation, stove monitoring, *etc.*). These needs are mostly addressed by a silo-based approach that requires a new assistive service (*e.g.*, a reminder system, a pill prompter) to be acquired for every activity to be supported. In practice, these services manifest their silo-based nature in their user interactions, and more specifically, in the heterogeneity of their notification system. This heterogeneity incurs a cognitive cost that prevents scaling up assistive services and compromises adoption by older adults. We present an approach to scaling up the combination of technology-based, assistive services by proposing a unifying notification system. To do so, (1) we propose a decomposition of assistive services to expose their needs in notification; (2) we introduce a notification framework, allowing heterogeneous assistive services to homogeneously notify users; (3) we present how this notification framework is carried out in practice for an assisted living platform. We successfully applied our approach to a range of existing and new assistive services. We used our notification framework to implement an assistive platform that combines a variety of assistive services. This platform has been deployed and used 24/7 in the home of 15 older adults for up to 6 months. This study provides empirical evidence of the effectiveness and learnability of the notification system of our platform, irrespective of the cognitive and sensory resources of the user. Additional results show that our assisted living platform achieved high user acceptance and satisfaction.

7.5. Orchestrating Masses of Sensors: A Design-Driven Development Approach

We propose a design-driven development approach that is dedicated to the domain of orchestration of masses of sensors. The developer declares what an application does using a domain-specific language (DSL). Our compiler processes domain-specific declarations to generate a customized programming framework that guides and supports the programming phase.

7.6. Analysis of How People with Intellectual Disabilities Organize Information Using Computerized Guidance

Access to residential settings for people with intellectual disabilities (ID) contributes to their social participation, but presents particular challenges. Assistive technologies can help people perform activities of daily living. However, the majority of the computerized solutions offered use guidance modes with a fixed, unchanging sequencing that leaves little room for self-determination to emerge. The objective of the project was to develop a flexible guidance mode and to test it with participants, to describe their information organization methods. This research used a descriptive exploratory design and conducted a comparison between five participants with ID and five participants with no ID. The results showed a difference in the information organization methods for both categories of participants. The people with ID used more diversified organization methods (categorical, schematic, action-directed) than the neurotypical participants (visual, action-directed). These organization methods varied depending on the people, but also on the characteristics of the requested task. Furthermore, several people with ID presented difficulties when switching from virtual to real mode. These results demonstrate the importance of developing flexible guidance modes adapted to the users' cognitive strategies, to maximize their benefits. Studies using experimental designs will have to be conducted to determine the impacts of more-flexible guidance modes.

RAP Project-Team

4. New Results

4.1. Random Graphs

Participant: Nicolas Broutin.

And/Or trees for random Boolean functions

For some time, a number of teams have tried to devise natural probability distributions on Boolean functions. Indeed, the most natural one, the uniform one, is not quite satisfactory: almost all Boolean functions have maximal complexity, while it is extremely difficult to construct some with high complexity. One approach consists in generating functions by seeing them as "expressions" encoded as a tree of computation. We generalize and unify the previous approaches that are restricted to very specific cases by looking at the distributions induced on the Boolean function by large computation trees that are arbitrary, except for the fact they the neighborhoods of the root (where the computation concentrates) stabilizes in distribution as the sizes of the tree increases [12].

4.2. Resource Allocation in Large Data Centres

Participants: Christine Fricker, Philippe Robert, Guilherme Thompson.

With the exponential increase in internet data transmission volume over the past years, efficient bandwidth allocation in large data centres has become crucial. Illustrating examples are the rapid spread of cloud computing technology, as well as the growth of the demand for video streaming, both of which were quasi non-existent 10 years ago.

Currently, most systems operate under decentralised policies due to the complexity of managing data exchange on large scales. In such systems, customer demands are served respecting their initial service requirements (a certain video quality, amount of memory or processing power etc.) until the system reaches saturation, which then leads to the blockage of subsequent customer demands. Strategies that rely on the scheduling of tasks are often not suitable to address this load balancing problem as the users expect instantaneous service usage in real time applications, such as video transmission and elastic computation. Our research goal is to understand and redesign its algorithms in order to develop decentralised policies that can improve global performance using local instantaneous information. This research is made in collaboration with Fabrice Guillemin, from Orange Labs.

In a first approach to this problem, we examined offloading schemes in fog computing context, where one data centres are installed at the edge of the network. We analyse the case with one data centre close to user which is backed up by a central (bigger) data centre. When a request arrives at an overloaded data centre, it is forwarded to the other data centre with a given probability, in order to help coping with saturation and reducing the rejection of requests. In [16], we have been able to show that the performance of such a system can be expressed in terms of the invariant distribution of a random walk in the quarter plane. As a consequence we have been able to assess the behaviour and performance of these systems, proving the effectiveness of such an offloading arrangement.

In a second step, we investigated allocation schemes which consist in reducing the bandwidth of arriving requests to a minimal value when the system is close to saturation. We analysed the effectiveness of such a downgrading policy, which, if the system is correctly designed, will reduce the fraction of rejected transmissions. We developed a mathematical model which allows us to predict system behaviour under such a policy and calculate the ideal threshold (in the same scale as the resource) after which downgrading should be initiated, given system parameters. We proved the existence of a unique equilibrium point, around which we have been able to determine the probability of the system being above or under the threshold. We found that system blockage can be almost surely eliminated. This policy finds a natural application in the context of video streaming services and other real time applications, such as MPEG-DASH. A document is being written to further publication.

Finally, with those results, we now try to extend our research towards more complex systems, investigating the behaviour of multiple resource systems (such as a Cloud environment, where computational power is provided using unities of CPU and GB of RAM) and other offloading schemes, such as the compulsory forwarding of a request when it's blocked at the edge server, but keeping a trunk reservation to protect the service originally assigned to the big data centre.

4.3. Resource allocation in vehicle sharing systems

Participants: Christine Fricker, Plinio Santini Dester, Hanene Mohamed, Yousra Chabchoub.

This is a collaboration with Danielle Tibi, Université Denis Diderot.

Vehicle sharing systems are becoming an urban mode of transportation, and launched in many cities, as Velib' and Autolib' in Paris. One of the major issues is the availability of the resources: vehicles or free slots to return them. These systems became a hot topic in Operation Research and now the importance of stochasticity on the system behavior is commonly admitted. The problem is to understand the system behavior and how to manage these systems in order to provide both resources to users. Our stochastic model is the first one taking into account the finite number of spots at the stations.

With Danielle Tibi, we use limit local theorems to obtain the asymptotic stationary joint distributions of several station states when the system is large (both numbers of stations and bikes), in the case of finite capacities of the stations. This gives an asymptotic independence property for node states. This widely extends the existing results on heterogeneous bike-sharing systems.

Recently we investigate some network load balancing algorithms to improve the bike sharing system behavior. We focus on the choice of the least loaded station among two to return the bike. A problem is the influence of the delay between the choice time (the beginning of the trip) and the time the station is joined (the end of the trip). However the main challenge is to deal with the choice between two neighboring stations. For that, a system of infinite queues is studied in light traffic. For a bike-sharing homogeneous model, we restrict our study to a deterministic cooperation of two by two stations. It relies on new results for the classical system of two queues under the join-the-shortest-queue policy.

JC Decaux provides us data describing Velib' user trips. These data are useful to measure the system parameters, validate our models and test our algorithms. Indeed, we use these data to investigate load balancing algorithms such as two-choice policies.

4.4. Scaling Methods

Participants: Philippe Robert, Wen Sun.

4.4.1. Fluid Limits in Wireless Networks

This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog: if a node of this network has x requests to transmit, then it receives a fraction of the capacity proportional to $\log(1+x)$, the logarithm of its current load. This year we completed the analysis of a star network topology with multiple nodes. Several scalings were used to describe the fluid limit behaviour.

4.4.2. The Time Scales of a Transient Network

A large distributed system where users' files are duplicated on unreliable data servers is investigated. Due to a server breakdown, a copy of a file can be lost, it can be retrieved if another copy of the same file is stored on other servers. In the case where no other copy of a given file is present in the network, it is definitely lost. In order to have multiple copies of a given file, it is assumed that each server can devote a fraction of its processing capacity to duplicate files on other servers to enhance the durability of the system.

A trade-off is necessary between the bandwidth and the memory used for this back-up mechanism and the data loss rate. Back-up mechanisms already exist and have been studied thanks to simulation. To our knowledge, no theoretical study exists on this topic. With a very simple centralized model, we have been able to emphasise a trade-off between capacity and life-time with respect to the duplication rate. From a mathematical point of view, we are currently studying different time scales of the system with an averaging phenomenon.

We have used scaling methods with different time scales to derive some asymptotic results on the decay of a simplified network: it is assumed that any copy of a given file is lost at some fixed rate and the total processing capacity of the system is devoted to duplicate the file with least number of copies. We start from the optimal initial state: each file has the maximum number of copies. Due to random losses, the state of the network is transient and all files will be eventually lost. There is a stability assumption for the system having a critical time scale of decay. When the stability condition is not satisfied, i.e. when it is initially overloaded, we have shown that the state of the network converges to an interesting local equilibrium. We are currently studying a more general case which the duplication depends on the structure of the system. See [7].

4.5. Stochastic Models of Biological Networks

Participants: Renaud Dessalles, Sarah Eugene, Philippe Robert.

4.5.1. Stochastic Modelling of self-regulation in the protein production system of bacteria

This is a collaboration with Vincent Fromion from INRA Jouy-en-Josas, which started on December 2014.

In prokaryotic cells (e.g. E. Coli. or B. Subtilis) the protein production system has to produce in a cell cycle (i.e. less than one hour) more than 10^6 molecules of more than 2500 kinds, each having different level of expression. The bacteria uses more than 85% of its resources to the protein production. Gene expression is a highly stochastic process: bacteria sharing the same genome, in a same environment will not produce exactly the same amount of a given protein. Some of this stochasticity can be due to the system of production itself: molecules, that take part in the production process, move freely into the cytoplasm and therefore reach any target in the cell after some random time; some of them are present in so much limited amount that none of them can be available for a certain time; the gene can be deactivated by repressors for a certain time, etc. We study the integration of several mechanisms of regulation and their performances in terms of variance and distribution. As all molecules tends to move freely into the cytoplasm, it is assumed that the encounter time between a given entity and its target is exponentially distributed.

4.5.1.1. Feedback model

We have also investigated the production of a single protein, with the transcription and the translation steps, but we also introduced a direct feedback on it: the protein tends to bind on the promoter of its own gene, blocking therefore the transcription. The protein remains on it during an exponential time until its detachment caused by thermal agitation.

The mathematical analysis aims at understanding the nature of the internal noise of the system and to quantify it. We tend to test the hypothesis usually made that such feedback permits a noise reduction of protein distribution compared to the "open loop" model. We have made the mathematical analysis of the model (using a scaling to be able to have explicit results), it appeared that reduction of variance compared to an "open loop" model is limited: the variance cannot be reduced for more than 50%.

We proposed another possible effect of the feedback loop: the return to equilibrium is faster in the case of a feedback model compared to the open loop model. Such behaviour can be beneficial for the bacteria to change of command for a new level of production of a particular protein (due, for example, to a radical change in the environment) by reducing the respond time to reach this new average. This study has been mainly performed by simulation and it has been shown that the feedback model can go 50% faster than the open loop results. See [13].

4.5.1.2. Transcription-translation model for all proteins

The other model that has been studied integrates the production of all the proteins. Each gene has to be transcribed in mRNA (using RNA-Polymerase molecules) and each mRNA has to be translated in protein (using ribosome molecules). Experiments (as the one from Taniguchi et al. (2010)) have shown that protein production is subject to high variability especially for highly expressed proteins. Our goal is to determine what in the protein production mechanism is responsible for the noise.

We already made simulations that takes into amount of RNA-Polymerases and Ribosomes and that genes and mRNAs sequester these molecules during the whole the time of elongation. This global sharing of Ribosomes/RNA-Polymerases reproduce only a part of the unknown noise experimentally seen. We are developing Python simulations that extends this model and take into account other feature that might be responsible for the noise in protein production. This new simulation will include new features such as:

- The volume of the cell. We consider it as proportional to the total number of proteins, and will increase as the cell grows. Transcription and translation initiation are then depending on the concentration of respectively free RNA-polymerase and free ribosomes.
- The division of the cell. At division, all components have an equal chance to go in either one of the two daughter cell.
- DNA replication. At some point in the cell cycle, the genome duplicates, doubling therefore the copy number of each gene

The simulation parameters will be fit with the data of Taniguchi et al. (2010) and the goal is to compare our result to see if which aspects of the protein production are responsible for the noise of the proteins.

4.5.2. Stochastic Modelling of Protein Polymerization

This is a collaboration with Marie Doumic, Inria MAMBA team.

The first part of our work focuses on the study of the polymerization of protein. This phenomenon is involved in many neurodegenerative diseases such as Alzheimer's and Prion diseases, e.g. mad cow. In this context, it consists in the abnormal aggregation of proteins. Curves obtained by measuring the quantity of polymers formed in in vitro experiments are sigmoids: a long lag phase with almost no polymers followed by a fast consumption of all monomers. Furthermore, repeating the experiment under the same initial conditions leads to somewhat identical curves up to translation. After having proposed a simple model to explain this fluctuations, we studied a more sophisticated model, closer to the reality. We added a conformation step: before being able to polymere, proteins have to misfold. This step is very quick and remains at equilibrium during the whole process. Nevertheless, this equilibrium depends on the polymerization which is happening on a slower time scale. The analysis of these models involves stochastic averaging principles.

The second part concerns the study of telomeres. This work is made in collaboration with Zhou Xu, Teresa Teixeira, from IBCP in Paris.

In eukaryotic cells, at each mitosis, chromosomes are shortened, because the DNA polymerase is not able to duplicate one ending of the chromosome. To prevent loss of genetic information- which could be catastrophic for the cell- chromosomes are equipped with telomeres at their endings. These telomeres do not contain any genetic information; they are a repetition of the sequence T-T-A-G-G-G thousands times. At each mitosis, there is therefore a loss of telomere. As it has a finite length, when the telomeres are too short, the cell cannot divide anymore: they enter in replicative senescence. Our model tries to captures the two phases of the shortening of telomeres: first, the initial state of the cells, when the telomerase is still active to repair the telomeres. Second, when the telomerase is inhibited, we try to estimate the senescence threshold, when the replication of the cells stops.

REGAL Project-Team

6. New Results

6.1. Distributed algorithms for dynamic networks

Participants: Luciana Bezerra Arantes [correspondent], Marjorie Bournat, Swan Dubois, Denis Jeanneau, Mohamed Hamza Kaaouachi, Sébastien Monnet, Franck Petit [correspondent], Pierre Sens, Julien Sopena.

Nowadays, distributed systems are more and more heterogeneous and versatile. Computing units can join, leave or move inside a global infrastructure. These features require the implementation of dynamic systems, that is to say they can cope autonomously with changes in their structure in terms of physical facilities and software. It therefore becomes necessary to define, develop, and validate distributed algorithms able to managed such dynamic and large scale systems, for instance mobile *ad hoc* networks, (mobile) sensor networks, P2P systems, Cloud environments, robot networks, to quote only a few.

We have obtained results both on fundamental aspects of distributed algorithms and on specific emerging large-scale applications.

We study various key topics of distributed algorithms: agreement, failure detection, data dissemination and data finding in large scale systems, self-stabilization and self-* services.

6.1.1. Agreement and failure detection in dynamic Distributed Systems

Distributed systems should provide reliable and continuous services despite the failures of some of their components. A classical way for a distributed system to tolerate failures is to detect them and then to recover. It is now well recognized that the dominant factor in system unavailability lies in the failure detection phase. In 2015, we obtain the following results on failure detection:

Assuming a message-passing environment with a majority of correct processes, the necessary and sufficient information about failures for implementing a general state machine replication scheme ensuring consistency is captured by the Ω failure detector. We show in [46] that in such a message-passing environment, Ω is also the weakest failure detector to implement an eventually consistent replicated service, where replicas are expected to agree on the evolution of the service state only after some (a priori unknown) time.

We also study the k-set agreement problem is a generalization of the consensus problem where processes can decide up to k different values. Very few papers have tackled this problem in dynamic networks. Exploiting the formalism of the Time Varying Graph model, we propose in [70] a new quorum-based failure detector for solving k-set agreement in dynamic networks with asynchronous communications. We present two algorithms that implement this new failure detector using graph connectivity and message pattern assumptions. We also provide an algorithm for solving k-set agreement using our new failure detector.

We propose several algorithms to implement efficient failure detection services. We introduce in [60] the Two Windows Failure Detector (2WFD), an algorithm that provides QoS and is able to react to sudden changes in network conditions, a property that currently existing algorithms do not satisfy. We ran tests on real traces and compared the 2W-FD to state-of-the-art algorithms. Our results show that our algorithm presents the best performance in terms of speed and accuracy in unstable scenarios. In [62], we propose a new approach towards the implementation of failure detectors for large and dynamic networks: we study reputation systems as a means to detect failures. The reputation mechanism allows efficient node cooperation via the sharing of views about other nodes. Our experimental results show that a simple prototype of a reputation-based detection service performs better than other known adaptive failure detectors, with improved flexibility. It can thus be used in a dynamic environment with a large and variable number of nodes.

6.1.2. Probabilistic Byzantine Tolerance allocation strategies in Hybrid Cloud Environments

We explore the node allocation challenges in providing probabilistic Byzantine fault tolerance in a hybrid cloud environment, consisting of nodes with varying reliability levels, compute power, and monetary cost. We consider hybrid computing architectures that combine edge nodes with cloud hosted computing. In such a system, a large fraction of the computation is performed by donated machines at the edge of the network, which significantly reduces the cost to the owner of the computation.

Considering “bag of tasks” (BoT) applications where a large computational problem is broken into a large number of independent tasks, the probabilistic Byzantine fault tolerance guarantee refers to the confidence level that the result of a given computation is correct despite potential Byzantine failures. In [36] we explore probabilistic Byzantine tolerance, in which computation tasks are replicated on dynamic replication sets whose size is determined based on ensuring probabilistic thresholds of correctness.

6.1.3. Covering problems in dynamic systems

We study covering problems (such as minimal dominating set or maximal matching) in the context of highly dynamic distributed systems. We first obtain some general results. In [48], we first propose a new definition of this family of problems since classical ones are meaningless in such systems. We generalize the classical definition of time complexity (for static systems) to our setting. We also provided in [40] a generic tool to help the writing of impossibility proofs in dynamic distributed systems. Then, we focus on the particular case of the minimal dominating set problem. We characterize the necessary and sufficient condition to construct deterministically a minimal dominating set in a dynamic system according to our definition.

6.1.4. Self-Stabilization

Self-stabilization is a generic paradigm to tolerate transient faults (*i.e.*, faults of finite duration) in distributed systems. Results obtained in this area by Regal members in 2015 follow.

Spanning tree construction is a well-studied problem in distributed computing for its numerous applications like routing, broadcast...Properties of the obtained trees, efficiency of the construction, and fault-tolerance guarantees are naturally at the heart of many researches. In this context, we propose in [39] a new self-stabilizing algorithm for the minimum diameter spanning tree that achieves better time and space complexity than existing solutions. Moreover, our solution tolerates a fully asynchronous adversary.

A classical way to endowed self-stabilization with (permanent) fault tolerance is *confinement*. That is, we ensure that the self-stabilizing system moreover ensures that the effect of permanent faults is limited to some topological areas of the system. In [27], we propose a characterization of optimal confinement areas for a large set of spanning tree metrics in presence of Byzantine faults. In [24], we propose a stabilizing implementation of an atomic register in presence of crash faults. By avoiding the propagation of fault effects further than a given radius, confinement is clearly a *spatial* approach. Another approach, called *temporal*, consists in recovering as quick as possible to a configuration from which some forms of safety are satisfied.

In [68], we introduce the notion of *gradual stabilization* and provide a gradually self-stabilizing algorithm that solves the *unison* problem, *i.e.*, the problem that consists in synchronizing logical clocks locally maintained by the processes.

6.1.5. Team of Mobile Robots

Swarm of autonomous mobile sensor devices (or, robots) recently emerged as an attractive issue in the study of dynamic distributed systems permits to assess the intrinsic difficulties of many fundamentals tasks, such as exploring or gathering in a discrete space. We consider autonomous robots that are endowed with visibility sensors (but that are otherwise unable to communicate) and motion actuators. The robots we consider are weak, *i.e.*, they are anonymous, uniform, unable to explicitly communicate, and oblivious (they do not remember any of their past actions). Despite their weakness, those robots must collaborate to solve a collective tasks such as exploration, gathering, flocking, to quote only a few.

In [45], we first show that it is impossible to explore any simple torus of arbitrary size with (strictly) less than four robots, even if the algorithm is probabilistic. Next, we propose an optimal (*w.r.t.* the number of robots) solution for the terminating exploration of torus-shaped networks by a team of k such robots in the SSYNC model. The proposed algorithm is probabilistic and works for any simple torus of size $\ell \times L$, where $7 \leq \ell \leq L$. Since the optimal number of robots is also four in rings, our result shows that increasing the number of possible symmetries in the network (due to increasing dimensions) does not necessarily come at an extra cost *w.r.t.* the number of robots that are necessary to solve the problem.

6.2. Management of distributed data

Participants: Rudyar Cortes, Mesaac Makpangou, Olivier Marin, Sébastien Monnet [correspondent], Pierre Sens.

6.2.1. Long term durability and storage load distribution

In 2014, we had proposed SPLAD (for Scattering and PLacing Data replicas to enhance long-term durability), a model that allows us to vary the data scattering degree by tuning a selection range width. We have enhanced our model [57] and we have focused on the study of the policy used while choosing a storing node within the selection range. Some policies may lead to heavily unbalanced storage load distribution which can be harmful for the system. Simple policies to balance the load (e.g. storing new blocks on least loaded nodes) may induce network congestion and thus data losses. We have shown that the “power of two choices” policy (choosing the least loaded node among two random ones) brings good results both in terms of storage load distribution and fault tolerance.

6.2.2. Management of dynamic big data

Managing and processing Dynamic Big Data, where multiple sources produce new data continuously, is very complex. Static cluster- or grid-based solutions are prone to induce bottleneck problems, and are therefore ill-suited in this context. Our objective in this domain is to design and implement a Reliable Large Scale Distributed Framework for the Management and Processing of Dynamic Big Data. In 2015, we focused on Spatio-temporal range queries over Big Location Data aim to extract and analyze relevant data items generated around a given location and time. They require concurrent processing of massive and dynamic data flows. We proposed a scalable architecture for continuous spatio-temporal range queries built by coalescing multiple computing nodes on top of a Distributed Hash Table. The key component of our architecture is a distributed spatio-temporal indexing structure which exhibits low insertion and low index maintenance costs. We assessed our solution with a public data set released by Yahoo! which comprises millions of geotagged multimedia files [43].

6.3. CISE Logic and tool for proving invariants in distributed databases

Participants: Marc Shapiro [correspondent], Mahsa Najafzadeh, Alexey Gotsman, Carla Ferreira.

We have developed a new sound logic for proving the correctness of a distributed database under concurrent updates, showing whether the application maintains the database’s *integrity invariants*. An operation of the application is specified as a *preparator*, which checks the operation’s precondition at an origin replica and generates an *effector*. The effector abstracts the update to be applied to every replica. The application also specifies which operations are allowed to take place concurrently. In summary, the logic shows that the application maintains the invariant if the three following rules are satisfied:

- Each operation individually maintains the invariant. It follows that operations’ preconditions are sufficiently strong to ensure correctness in a sequential execution.
- The effectors of any two operations that can execute concurrently commute. This implies that the database replicas all converge to the same state.
- For any pair of operations u and v that can execute concurrently, the precondition of u is stable under the effector of v , and vice-versa.

This result is published at POPL 2016 [50].

We have implemented a tool (based on the Z3 SMT solver) that implements these rules. A demo of the tool is available online [78]. If the application passes the tool, it is correct. If not, the tool returns a counter-example, which the application developer can inspect to find the source of the error. Generally speaking, the developer can either weaken the invariants or the effects of operations, or strengthen consistency by disallowing concurrency. By choosing one or the other, the developer performs a co-design of the application with its consistency protocol, in order to have the highest possible concurrency that still ensures correctness.

For instance, consider a database of bank accounts, with the invariant that an account's balance must be positive. The banking application has operations $credit(acct, amt)$, $debit(acct, amt)$, and $accrue - interest(acct)$. The first rule dictates that $debit$ has the precondition $amt = balance$. The second rule dictates that $accrue - interest$ computes the amount of interest according to the state at the origin, not at every replica. The third rule is violated if concurrent $debits$ are allowed; if the bank wishes to uphold the invariant, the only correct solution is to disallow concurrent $debits$.

6.4. Memory management for big data

Participants: Antoine Blin, Damien Carver, Maxime Lorrillere, Sébastien Monnet, Julien Sopena [correspondent].

6.4.1. Automated file cache pooling

Some applications, like online sales servers, intensively use disk I/Os. Their performance is tightly coupled with I/Os efficiency. To speed up I/Os, operating systems use free memory to offer caching mechanisms. Several I/O intensive applications may require a large cache to perform well. However, nowadays resources are virtualized. In clouds, for instance, virtual machines (VMs) offer both isolation and flexibility. This is the foundation of cloud elasticity, but it induces fragmentation of the physical resources, including memory. This fragmentation reduces the amount of available memory a VM can use for caching I/Os. Previously, we proposed Puma (for Pooling Unused Memory in Virtual Machines) which allows I/O intensive applications running on top of VMs to benefit of large caches. This was realized by providing a remote caching mechanism that provides the ability for any VM to extend its cache using the memory of other VMs located either in the same or in a different host.

We have performed an extensive evaluation of Puma [53] and we have enhanced our solution: Puma adapts automatically the amount a memory that a VM offers to another VM. Furthermore, if the network becomes overloaded, Puma detects a performance degradation and stops using a remote cache.

RMOD Project-Team

7. New Results

7.1. Tools for understanding evolution

Automatic Detection of System-Specific Conventions. In Apache Ant, a convention to improve maintenance was introduced in 2004 stating a new way to close files instead of the Java generic `InputStream.close()`. Yet, six years after its introduction, this convention was still not generally known to the developers. Two existing solutions could help in these cases. First, one can deprecate entities, but, in our example, one can hardly deprecate Java's method. Second, one can create a system-specific rule to be automatically enforced. In a preceding publication, we showed that system-specific rules are more likely to be noticed by developers than generic ones. However, in practice, developers rarely create specific rules. We therefore propose to free the developers from the need to create rules by automatically detecting such conventions from source code repositories. This is done by mining the change history of the system to discover similar changes being applied over several revisions. The proposed approach is applied to real-world systems, and the extracted rules are validated with the help of experts. The results show that many rules are in fact relevant for the experts. [16]

DeltaImpactFinder. In software development, version control systems (VCS) provide branching and merging support tools. Such tools are popular among developers to concurrently change a code-base in separate lines and reconcile their changes automatically afterwards. However, two changes that are correct independently can introduce bugs when merged together. We call semantic merge conflicts this kind of bugs. Change impact analysis (CIA) aims at estimating the effects of a change in a codebase. We propose to detect semantic merge conflicts using CIA. On a merge, DELTAIMPACTFINDER analyzes and compares the impact of a change in its origin and destination branches. We call the difference between these two impacts the delta-impact. If the delta-impact is empty, then there is no indicator of a semantic merge conflict and the merge can continue automatically. Otherwise, the delta-impact contains what are the sources of possible conflicts. [26]

OrionPlanning. Many techniques have been proposed in the literature to support architecture definition, conformance, and analysis. However, there is a lack of adoption of such techniques by the industry. Previous work have analyzed this poor support. Specifically, former approaches lack proper analysis techniques (e.g., detection of architectural inconsistencies), and they do not provide extension and addition of new features. We present ORIONPLANNING, a prototype tool to assist refactorings at large scale. The tool provides support for model-based refactoring operations. These operations are performed in an interactive visualization. The contributions of the tool consist in: (i) providing iterative modifications in the architecture, and (ii) providing an environment for architecture inspection and definition of dependency rules. [37]

Recording and Replaying System-Specific Conventions. During its lifetime, a software system is under continuous maintenance to remain useful. Maintenance can be achieved in activities such as adding new features, fixing bugs, improving the system's structure, or adapting to new APIs. In such cases, developers sometimes perform sequences of code changes in a systematic way. These sequences consist of small code changes (e.g., create a class, then extract a method to this class), which are applied to groups of related code entities (e.g., some of the methods of a class). MacroRecorder is a proof-of-concept tool that records a sequence of code changes, then it allows the developer to generalize this sequence in order to apply it in other code locations. The evaluation is based on previous work on repetitive code changes related to rearchitecting. MacroRecorder was able to replay 92% of the examples, which consisted in up to seven code entities modified up to 66 times. The generation of a customizable, large-scale transformation operator has the potential to efficiently assist code maintenance. [39], [38]

7.2. Software Quality: Taming Software Evolution

Software metrics do not predict the health of a project. More and more companies would like to mine software data with the goal of assessing the health of their software projects. The hope is that some software

metrics could be tracked to predict failure risks or confirm good health. If a factor of success was found, projects failures could be anticipated and early actions could be taken by the organisation to help or to monitor closely the project, allowing one to act in a preventive mode rather than a curative one. We were called by a major IT company to fulfil this goal. We conducted a study to check whether software metrics can be related to project failure. The study was both theoretic with a review of literature on the subject, and practical with mining past projects data and interviews with project managers. We found that metrics used in practice are not reliable to assess project outcome. [22]

How Do Developers React to API Evolution? Software engineering research now considers that no system is an island, but it is part of an ecosystem involving other systems, developers, users, hardware, .. When one system (e.g., a framework) evolves, its clients often need to adapt. Client developers might need to adapt to functionalities, client systems might need to be adapted to a new API, client users might need to adapt to a new User Interface. The consequences of such changes are yet unclear, what proportion of the ecosystem might be expected to react, how long might it take for a change to diffuse in the ecosystem, do all clients react in the same way? We report on an exploratory study aimed at observing API evolution and its impact on a large-scale software ecosystem, Pharo, which has about 3,600 distinct systems, more than 2,800 contributors, and six years of evolution. We analyze 118 API changes and answer research questions regarding the magnitude, duration, extension, and consistency of such changes in the ecosystem. The results of this study help to characterize the impact of API evolution in large software ecosystems, and provide the basis to better understand how such impact can be alleviated. [27]

Does JavaScript software embrace classes? JavaScript is the de facto programming language for the Web. It is used to implement mail clients, office applications, or IDEs, that can weight hundreds of thousands of lines of code. The language itself is prototype based, but to master the complexity of their application, practitioners commonly rely on some informal class abstractions. This practice has never been the target of empirical investigations in JavaScript. Yet, understanding it would be key to adequately tune programming environments and structure libraries such as they are accessible to programmers. We report a large and in-depth study to understand how class emulation is employed in JavaScript applications. We propose a strategy to statically detect class-based abstractions in the source code of JavaScript systems. We used this strategy in a dataset of 50 popular JavaScript applications available from GitHub. We found systems structured around hundreds of classes, suggesting that JavaScript developers are standing on traditional class-based abstractions to tackle the growing complexity of their systems. [28]

7.3. Software Quality: History and Changes

Mining Architectural Violations from Version History. Software architecture conformance is a key software quality control activity that aims to reveal the progressive gap normally observed between concrete and planned software architectures. However, formally specifying an architecture can be difficult, as it must be done by an expert of the system having a high level understanding of it. We present a lightweight approach for architecture conformance based on a combination of static and historical source code analysis. The proposed approach relies on four heuristics for detecting absences (something expected was not found) and divergences (something prohibited was found) in source code based architectures. We also present an architecture conformance process based on the proposed approach. We followed this process to evaluate the architecture of two industrial-strength information systems, achieving an overall precision of 62.7% and 53.8%. We also evaluated our approach in an open-source information retrieval library, achieving an overall precision of 59.2%. We envision that an heuristic-based approach for architecture conformance can be used to rapidly raise architectural warnings, without deeply involving experts in the process. [17]

Untangling Fine-Grained Code Changes. After working for some time, developers commit their code changes to a version control system. When doing so, they often bundle unrelated changes (e.g., bug fix and refactoring) in a single commit, thus creating a so-called tangled commit. Sharing tangled commits is problematic because it makes review, reversion, and integration of these commits harder and historical analyses of the project less reliable. Researchers have worked at untangling existing commits, i.e., finding which part of a commit relates to which task. We contribute to this line of work in two ways: (1) A publicly available

dataset of untangled code changes, created with the help of two developers who accurately split their code changes into self contained tasks over a period of four months; (2) a novel approach, EpiceaUntangler, to help developers share untangled commits (aka. atomic commits) by using fine-grained code change information. EpiceaUntangler is based and tested on the publicly available dataset, and further evaluated by deploying it to 7 developers, who used it for 2 weeks. We recorded a median success rate of 91% and average one of 75%, in automatically creating clusters of untangled fine-grained code changes. [25]

Developers' Perception of Co-Change Patterns: An Empirical Study. Co-change clusters are groups of classes that frequently change together. They are proposed as an alternative modular view, which can be used to assess the traditional decomposition of systems in packages. To investigate developer's perception of co-change clusters, we report a study with experts on six systems, implemented in two languages. We mine 102 co-change clusters from the version history of such systems, which are classified in three patterns regarding their projection to the package structure: Encapsulated, Crosscutting, and Octopus. We then collect the perception of expert developers on such clusters, aiming to ask two central questions: (a) what concerns and changes are captured by the extracted clusters? (b) do the extracted clusters reveal design anomalies? We conclude that Encapsulated Clusters are often viewed as healthy designs and that Crosscutting Clusters tend to be associated to design anomalies. Octopus Clusters are normally associated to expected class distributions, which are not easy to implement in an encapsulated way, according to the interviewed developers. [40]

7.4. Dynamic Languages: Debugging

Practical domain-specific debuggers. Understanding the run-time behavior of software systems can be a challenging activity. Debuggers are an essential category of tools used for this purpose as they give developers direct access to the running systems. Nevertheless, traditional debuggers rely on generic mechanisms to introspect and interact with the running systems, while developers reason about and formulate domain-specific questions using concepts and abstractions from their application domains. This mismatch creates an abstraction gap between the debugging needs and the debugging support leading to an inefficient and error-prone debugging effort, as developers need to recover concrete domain concepts using generic mechanisms. To reduce this gap, and increase the efficiency of the debugging process, we propose a framework for developing domain-specific debuggers, called the Moldable Debugger, that enables debugging at the level of the application domain. The Moldable Debugger is adapted to a domain by creating and combining domain-specific debugging operations with domain-specific debugging views, and adapts itself to a domain by selecting, at run time, appropriate debugging operations and views. To ensure the proposed model has practical applicability (i.e., can be used in practice to build real debuggers), we discuss, from both a performance and usability point of view, three implementation strategies. We further motivate the need for domain-specific debugging, identify a set of key requirements and show how our approach improves debugging by adapting the debugger to several domains. [14]

Mercury: Properties and Design of a Remote Debugging Solution using Reflection. Remote debugging facilities are a technical necessity for devices that lack appropriate input/output interfaces (display, keyboard, mouse) for programming (e.g., smartphones, mobile robots) or are simply unreachable for local development (e.g., cloud-servers). Yet remote debugging solutions can prove awkward to use due to re-deployments. Empirical studies show us that on average 10.5 minutes per coding hour (over five 40-hour work weeks per year) are spent for redeploying applications (including re-deployments during debugging). Moreover current solutions lack facilities that would otherwise be available in a local setting because it is difficult to reproduce them remotely. Our work identifies three desirable properties that a remote debugging solution should exhibit, namely: run-time evolution, semantic instrumentation and adaptable distribution. Given these properties we propose and validate Mercury, a remote debugging model based on reflection. Mercury supports run-time evolution through a causally connected remote meta-level, semantic instrumentation through the reification of the underlying execution environment and adaptable distribution through a modular architecture of the debugging middleware. [19]

7.5. Reconciling Dynamic Languages and Isolation

Handles. Controlling object graphs and giving specific semantics to references (such as read-only, ownership, scoped sharing) have been the focus of a large body of research in the context of static type systems. Controlling references to single objects and to graphs of objects is essential to build more secure systems, but is notoriously hard to achieve in the absence of static type systems. In this article we embrace this challenge by proposing a solution to the following question: What is an underlying mechanism that can support the definition of properties (such as revocable, read-only, lent) at the reference level in the absence of a static type system? We present handles: first-class references that propagate behavioral change dynamically to the object subgraph during program execution. In this article we describe handles and show how handles support the implementation of read-only references and revocable references. Handles have been fully implemented by modifying an existing virtual machine and we report their costs. [13]

Delegation Proxies. Scoping behavioral variations to dynamic extents is useful to support non-functional concerns that otherwise result in cross-cutting code. Unfortunately, such forms of scoping are difficult to obtain with traditional reflection or aspects. We propose delegation proxies, a dynamic proxy model that supports behavioral intercession through the interception of various interpretation operations. Delegation proxies permit different behavioral variations to be easily composed together. We show how delegation proxies enable behavioral variations that can propagate to dynamic extents. We demonstrate our approach with examples of behavioral variations scoped to dynamic extents that help simplify code related to safety, reliability, and monitoring. [21]

Access Control to Reflection with Object Ownership. Reflection is a powerful programming language feature that enables language extensions, generic code, dynamic analyses, development tools, etc. However, uncontrolled reflection breaks object encapsulation and considerably increases the attack surface of programs e.g., malicious libraries can use reflection to attack their client applications. To bring reflection and object encapsulation back together, we use dynamic object ownership to design an access control policy to reflective operations. This policy grants objects full reflective power over the objects they own but limited reflective power over other objects. Code is still able to use advanced reflective operations but reflection cannot be used as an attack vector anymore. [41]

7.6. Tailoring Applications and bootstrapping

Virtualization Support for Dynamic Core Library Update. Dynamically updating language runtime and core libraries such as collections and threading is challenging since the update mechanism uses such libraries at the same time that it modifies them. To tackle this challenge, we present Dynamic Core Library Update (DCU) as an extension of Dynamic Software Update (DSU) and our approach based on a virtualization architecture. Our solution supports the update of core libraries as any other normal library, avoiding the circular dependencies between the updater and the core libraries. Our benchmarks show that there is no evident performance overhead in comparison with a default execution. Finally, we show that our approach can be applied to real life scenario by introducing a critical update inside a web application with 20 simulated concurrent users. [34]

Bootstrapping Infrastructure. Bootstrapping is well known in the context of compilers, where a bootstrapped compiler can compile its own source code. Bootstrapping is a beneficial engineering practice because it raises the level of abstraction of a program making it easier to understand, optimize, evolve, etc. Bootstrapping a reflective object-oriented language is however more challenging, as we need also to initialize the runtime of the language with its initial objects and classes besides writing its compiler. We present a novel bootstrapping infrastructure for Pharo-like languages that allows us to easily extend and modify such languages. Our bootstrapping process relies on a first-class runtime. A first-class runtime is a meta-object that represents a program's runtime and provides a MOP to easily load code into it and manipulate its objects. It decouples the virtual machine (VM) and language concerns by introducing a clear VM-language interface. Using this process, we show how we succeeded to bootstrap a Smalltalk-based language named Candle and then extend it with traits in less than 250 lines of high-level Smalltalk code. We also show how we can bootstrap with minimal effort two other languages (Pharo and MetaTalk) with similar execution semantics but different object models. [35]

7.7. Dynamic Languages: Virtual Machines

Towards Fully Reflective Environments. Modern development environments promote live programming (LP) mechanisms because it enhances the development experience by providing instantaneous feedback and interaction with live objects. LP is typically supported with advanced reflective techniques within dynamic languages. These languages run on top of Virtual Machines (VMs) that are built in a static manner so that most of their components are bound at compile time. As a consequence, VM developers are forced to work using the traditional edit-compile-run cycle, even when they are designing LP-supporting environments. We explore the idea of bringing LP techniques to VM development to improve the observability, evolution and adaptability of VMs at run-time. We define the notion of fully reflective execution environments, systems that provide reflection not only at the application level but also at the level of the execution environment (EE). We characterize such systems, propose a design, and present Mate v1, a prototypical implementation. Based on our prototype, we analyze the feasibility and applicability of incorporating reflective capabilities into different parts of EEs. Furthermore, the evaluation demonstrates the opportunities such reflective capabilities provide for unanticipated dynamic adaptation scenarios, benefiting thus, a wider range of users. [23]

Tracing vs. Partial Evaluation. Tracing and partial evaluation have been proposed as meta-compilation techniques for interpreters to make just-in-time compilation language-independent. They promise that programs executing on simple interpreters can reach performance of the same order of magnitude as if they would be executed on state-of-the-art virtual machines with highly optimizing just-in-time compilers built for a specific language. Tracing and partial evaluation approach this meta-compilation from two ends of a spectrum, resulting in different sets of tradeoffs. This study investigates both approaches in the context of self-optimizing interpreters, a technique for building fast abstract-syntax-tree interpreters. Based on RPython for tracing and Truffle for partial evaluation, we assess the two approaches by comparing the impact of various optimizations on the performance of an interpreter for SOM, an object-oriented dynamically-typed language. The goal is to determine whether either approach yields clear performance or engineering benefits. We find that tracing and partial evaluation both reach roughly the same level of performance. SOM based on meta-tracing is on average 3x slower than Java, while SOM based on partial evaluation is on average 2.3x slower than Java. With respect to the engineering, tracing has however significant benefits, because it requires language implementers to apply fewer optimizations to reach the same level of performance. [29]

Zero-Overhead Metaprogramming. Runtime metaprogramming enables many useful applications and is often a convenient solution to solve problems in a generic way, which makes it widely used in frameworks, middleware, and domain-specific languages. However, powerful metaobject protocols are rarely supported and even common concepts such as reflective method invocation or dynamic proxies are not optimized. Solutions proposed in literature either restrict the metaprogramming capabilities or require application or library developers to apply performance improving techniques. For overhead-free runtime metaprogramming, we demonstrate that dispatch chains, a generalized form of polymorphic inline caches common to self-optimizing interpreters, are a simple optimization at the language-implementation level. Our evaluation with self-optimizing interpreters shows that unrestricted metaobject protocols can be realized for the first time without runtime overhead, and that this optimization is applicable for just-in-time compilation of interpreters based on meta-tracing as well as partial evaluation. In this context, we also demonstrate that optimizing common reflective operations can lead to significant performance improvements for existing applications [30].

A Partial Read Barrier for Efficient Support of Live Object-oriented Programming. Live programming, originally introduced by Smalltalk and Lisp, and now gaining popularity in contemporary systems such as Swift, requires on-the-fly support for object schema migration, such that the layout of objects may be changed while the program is at one and the same time being run and developed. In Smalltalk schema migration is supported by two primitives, one that answers a collection of all instances of a class, and one that exchanges the identities of pairs of objects, called the become primitive. Existing instances are collected, copies using the new schema created, state copied from old to new, and the two exchanged with become, effecting the schema migration. Historically the implementation of become has either required an extra level of indirection between an object's address and its body, slowing down slot access, or has required a sweep of all objects, a very slow operation on large heaps. Spur, a new object representation and memory manager for Smalltalk-like

languages, has neither of these deficiencies. It uses direct pointers but still provides a fast become operation in large heaps, thanks to forwarding objects that when read conceptually answer another object and a partial read barrier that avoids the cost of explicitly checking for forwarding objects on the vast majority of object accesses [31].

ROMA Project-Team

7. New Results

7.1. Scheduling computational workflows on failure-prone platforms

Participants: Guillaume Aupy, Anne Benoit, Henri Casanova [University of Hawaii], Yves Robert.

We study the scheduling of computational workflows on compute resources that experience exponentially distributed failures. When a failure occurs, rollback and recovery is used to resume the execution from the last checkpointed state. The scheduling problem is to minimize the expected execution time by deciding in which order to execute the tasks in the workflow and whether to checkpoint or not checkpoint a task after it completes. We give a polynomial-time algorithm for fork graphs and show that the problem is NP-complete with join graphs. Our main result is a polynomial-time algorithm to compute the execution time of a workflow with specified to-be-checkpointed tasks. Using this algorithm as a basis, we propose efficient heuristics for solving the scheduling problem. We evaluate these heuristics for representative workflow configurations.

This work has been published in the 17th Workshop on Advances in Parallel and Distributed Computational Models [20].

7.2. Efficient checkpoint/verification patterns

Participants: Anne Benoit, Saurabh K. Raina [Jaypee Institute of Information Technology], Yves Robert.

Errors have become a critical problem for high performance computing. Checkpointing protocols are often used for error recovery after fail-stop failures. However, silent errors cannot be ignored, and their peculiarity is that such errors are identified only when the corrupted data is activated. To cope with silent errors, we need a verification mechanism to check whether the application state is correct. Checkpoints should be supplemented with verifications to detect silent errors. When a verification is successful, only the last checkpoint needs to be kept in memory because it is known to be correct.

In this work, we analytically determine the best balance of verifications and checkpoints so as to optimize platform throughput. We introduce a balanced algorithm using a pattern with p checkpoints and q verifications, which regularly interleaves both checkpoints and verifications across same-size computational chunks. We show how to compute the waste of an arbitrary pattern, and we prove that the balanced algorithm is optimal when the platform MTBF (Mean Time Between Failures) is large in front of the other parameters (checkpointing, verification and recovery costs). We conduct several simulations to show the gain achieved by this balanced algorithm for well-chosen values of p and q , compared to the base algorithm that always perform a verification just before taking a checkpoint ($p = q = 1$), and we exhibit gains of up to 19%.

This work has been published in the International Journal of High Performance Computing Applications [8].

7.3. Assessing the impact of partial verifications against silent data corruptions

Participants: Aurélien Cavelan, Saurabh K. Raina [Jaypee Institute of Information Technology], Yves Robert, Hongyang Sun.

Silent errors, or silent data corruptions, constitute a major threat on very large scale platforms. When a silent error strikes, it is not detected immediately but only after some delay, which prevents the use of pure periodic checkpointing approaches devised for fail-stop errors. Instead, checkpointing must be coupled with some verification mechanism to guarantee that corrupted data will never be written into the checkpoint file. Such a guaranteed verification mechanism typically incurs a high cost. In this work, we assess the impact of using partial verification mechanisms in addition to a guaranteed verification. The main objective is to investigate to which extent it is worthwhile to use some light cost but less accurate verifications in the middle of a periodic computing pattern, which ends with a guaranteed verification right before each checkpoint. Introducing partial verifications dramatically complicates the analysis, but we are able to analytically determine the optimal computing pattern (up to the first-order approximation), including the optimal length of the pattern, the optimal number of partial verifications, as well as their optimal positions inside the pattern. Performance evaluations based on a wide range of parameters confirm the benefit of using partial verifications under certain scenarios, when compared to the baseline algorithm that uses only guaranteed verifications.

This work has been published in the proceedings of ICCP'15 [22].

7.4. Which Verification for Soft Error Detection?

Participants: Leonardo Bautista-Gomez [Argonne National Laboratory], Anne Benoit, Aurélien Cavelan, Saurabh K. Raina [Jaypee Institute of Information Technology], Yves Robert, Hongyang Sun.

This work is an extension of the work described in Section 7.4 to cope with imperfect verifications. Many methods are available to detect silent errors in high-performance computing (HPC) applications. Each comes with a given cost and recall (fraction of all errors that are actually detected). The main contribution of this work is to characterize the optimal computational pattern for an application: which detector(s) to use, how many detectors of each type to use, together with the length of the work segment that precedes each of them. We conduct a comprehensive complexity analysis of this optimization problem, showing NP-completeness and designing an FPTAS (Fully Polynomial-Time Approximation Scheme). On the practical side, we provide a greedy algorithm whose performance is shown to be close to the optimal for a realistic set of evaluation scenarios.

This work has been published in the proceedings of HiPC'15 [21].

7.5. Composing resilience techniques: ABFT, periodic and incremental checkpointing

Participants: George Bosilca [University of Tennessee, Knoxville], Aurélien Bouteiller [University of Tennessee, Knoxville], Thomas Hérault [University of Tennessee, Knoxville], Yves Robert, Jack Dongarra [University of Tennessee, Knoxville].

Algorithm Based Fault Tolerant (ABFT) approaches promise unparalleled scalability and performance in failure-prone environments. Thanks to recent advances in the understanding of the involved mechanisms, a growing number of important algorithms (including all widely used factorizations) have been proven ABFT-capable. In the context of larger applications, these algorithms provide a temporal section of the execution, where the data is protected by its own intrinsic properties, and can therefore be algorithmically recomputed without the need of checkpoints. However, while typical scientific applications spend a significant fraction of their execution time in library calls that can be ABFT-protected, they interleave sections that are difficult or even impossible to protect with ABFT. As a consequence, the only practical fault-tolerance approach for these applications is checkpoint/restart. In this work, we propose a model to investigate the efficiency of a composite protocol, that alternates between ABFT and checkpoint/restart for the effective protection of an iterative application composed of ABFT-aware and ABFT-unaware sections. We also consider an incremental checkpointing composite approach in which the algorithmic knowledge is leveraged by a novel optimal dynamic programming to compute checkpoint dates. We validate these models using a simulator. The model and simulator show that the composite approach drastically increases the performance delivered by an execution platform, especially at scale, by providing the means to increase the interval between checkpoints while simultaneously decreasing the volume of each checkpoint.

This work has been published in the International Journal of Networking and Computing [9].

7.6. Voltage Overscaling Algorithms for Energy-Efficient Workflow Computations With Timing Errors

Participants: Aurélien Cavelan, Yves Robert, Hongyang Sun, Frédéric Vivien.

We proposed a software-based approach using dynamic voltage overscaling to reduce the energy consumption of HPC applications. This technique aggressively lowers the supply voltage below nominal voltage, which introduces timing errors, and we used Algorithm-Based Fault-Tolerance (ABFT) to provide fault tolerance for matrix operations. We introduced a formal model, and we designed optimal polynomial-time solutions, to execute a linear chain of tasks. Evaluation results obtained for matrix multiplication demonstrated that our approach indeed leads to significant energy savings, compared to the standard algorithm that always operates at nominal voltage.

This work has been published in the proceedings of the 5th Workshop on Fault Tolerance for HPC at eXtreme Scale [23].

7.7. Approximation algorithms for energy, reliability and makespan optimization problems

Participants: Guillaume Aupy, Anne Benoit.

We consider the problem of scheduling an application on a parallel computational platform. The application is a particular task graph, either a linear chain of tasks, or a set of independent tasks. The platform is made of identical processors, whose speed can be dynamically modified. It is also subject to failures: if a processor is slowed down to decrease the energy consumption, it has a higher chance to fail. Therefore, the scheduling problem requires us to re-execute or replicate tasks (i.e., execute twice the same task, either on the same processor, or on two distinct processors), in order to increase the reliability. It is a tri-criteria problem: the goal is to minimize the energy consumption, while enforcing a bound on the total execution time (the makespan), and a constraint on the reliability of each task.

Our main contribution is to propose approximation algorithms for linear chains of tasks and independent tasks. For linear chains, we design a fully polynomial-time approximation scheme. However, we show that there exists no constant factor approximation algorithm for independent tasks, unless $P=NP$, and we propose in this case an approximation algorithm with a relaxation on the makespan constraint.

This work has been published in the Parallel Processing Letters [4].

7.8. Co-scheduling algorithms for high-throughput workload execution

Participants: Guillaume Aupy, Manu Shantharam [San Diego Supercomputer Center], Anne Benoit, Yves Robert, Padma Raghavan [Penn State University].

This work investigates co-scheduling algorithms for processing a set of parallel applications. Instead of executing each application one by one, using a maximum degree of parallelism for each of them, we aim at scheduling several applications concurrently. We partition the original application set into a series of packs, which are executed one by one. A pack comprises several applications, each of them with an assigned number of processors, with the constraint that the total number of processors assigned within a pack does not exceed the maximum number of available processors. The objective is to determine a partition into packs, and an assignment of processors to applications, that minimize the sum of the execution times of the packs.

We thoroughly study the complexity of this optimization problem, and propose several heuristics that exhibit very good performance on a variety of workloads, whose application execution times model profiles of parallel scientific codes. We show that co-scheduling leads to faster workload completion time and to faster response times on average (hence increasing system throughput and saving energy), for significant benefits over traditional scheduling from both the user and system perspectives.

This work has been published in the Journal of Scheduling [6].

7.9. Scheduling the I/O of HPC Applications Under Congestion

Participants: Ana Gainaru [University of Illinois at Urbana Champaign], Guillaume Aupy, Anne Benoit, Franck Cappello, Yves Robert.

A significant percentage of the computing capacity of large-scale platforms is wasted due to interferences incurred by multiple applications that access a shared parallel file system concurrently. One solution to handling I/O bursts in large-scale HPC systems is to absorb them at an intermediate storage layer consisting of burst buffers. However, our analysis of the Argonne's Mira system shows that burst buffers cannot prevent congestion at all times. As a consequence, I/O performance is dramatically degraded, showing in some cases a decrease in I/O throughput of 67%.

In this work, we analyze the effects of interference on application I/O bandwidth, and propose several scheduling techniques to mitigate congestion. We focus on typical HPC applications, which have a periodic pattern consisting of some amount of computation followed by some volume of I/O to be transferred. We show through extensive experiments that our global I/O scheduler is able to reduce the effects of congestion, even on systems where burst buffers are used, and can increase the overall system throughput up to 56%. We also show that it outperforms current Mira I/O schedulers, even for non-periodic applications.

This work has been published in IPDPS'15 [26].

7.10. Scheduling trees of malleable tasks for sparse linear algebra

Participants: Abdou Guermouche [Univ. Bordeaux/Inria Bordeaux Sud-Ouest], Loris Marchal, Bertrand Simon, Oliver Sinnen [Univ. Auckland/New Zealand], Frédéric Vivien.

Scientific workloads are often described by directed acyclic task graphs. This is in particular the case for multifrontal factorization of sparse matrices —the focus of this work— whose task graph is structured as a tree of parallel tasks. Prasanna and Musicus [84], [85] advocated using the concept of *malleable* tasks to model parallel tasks involved in matrix computations. In this powerful model each task is processed on a time-varying number of processors. Following Prasanna and Musicus, we consider malleable tasks whose speedup is p^α , where p is the fractional share of processors on which a task executes, and α ($0 < \alpha \leq 1$) is a task-independent parameter. Firstly, we use actual experiments on multicore platforms to motivate the relevance of this model for our application. Then, we study the optimal time-minimizing allocation proposed by Prasanna and Musicus using optimal control theory. We greatly simplify their proofs by resorting only to pure scheduling arguments. Building on the insight gained thanks to these new proofs, we extend the study to distributed (homogeneous or heterogeneous) multicore platforms. We prove the NP-completeness of the corresponding scheduling problem, and we then propose some approximation algorithms [28].

In a second step, we studied a simplified speed-up function for malleable tasks, corresponding to perfect parallelism for a number of processors below a given threshold. The threshold depends on the task. We proved that scheduling independent chains of malleable tasks under this model is NP-complete. We study the performance of a classical allocation policy which is agnostic of the threshold and a simple greedy heuristic, and proved that both are 2-approximation algorithms, even if in practice, the latter often outperforms the former.

7.11. Parallel scheduling of task trees with limited memory

Participants: Clément Brasseur [ENS Lyon], Guillaume Aupy, Loris Marchal.

Scientific workloads are often described by directed acyclic task graphs. This is in particular the case for multifrontal factorization of sparse matrices —the focus of this work— whose task graph is structured as a tree of parallel tasks. When processing this tree on a multicore machine, we have to find a tradeoff between task parallelism and memory usage. In this context, Agullo et al. [62] proposed an activation scheme which follows a postorder traversal and books the memory needed for the task. This strategy has a low complexity and thus has been implemented in the lightweight runtime system StarPU [65], but may lead to excessive memory booking, which limits the task parallelism. In this work, we proposed a new booking strategy that books exactly what is necessary for a task, given what is already booked by its predecessors in the tree. We have shown by extensive simulations on realistic trees that this leads to better task parallelism and reduces the overall processing time.

7.12. Locality of Map tasks in MapReduce computations

Participants: Olivier Beaumont [Inria Bordeaux Sud-Ouest], Loris Marchal.

In data parallel system such as MapReduce, large data files are distributed among the storage attached to computing nodes, and the computation is afterwards allocated close to the data whenever it is possible. Several parameters may affect the locality of the data, and thus the amount of data that needs to be communicated during the computation: the possible replication of the data when it is distributed on the platform, and the load-balancing mechanism that transmits new data to node which have exhausted their own data. In this work, we have proposed a simple analytical model to estimate the amount of data transfer of various scenarios for the Map phase of MapReduce computations and we have validated this model using simulations.

7.13. Improving multifrontal methods by means of block low-rank representations

Participants: Patrick Amestoy [INPT-IRIT, Université of Toulouse], Cleve Ashcraft [LSTC], Olivier Boiteau [EDF], Alfredo Buttari [CNRS-IRIT, Université of Toulouse], Jean-Yves L'Excellent, Clément Weisbecker [INPT-IRIT, now at LSTC].

Matrices coming from elliptic Partial Differential Equations (PDEs) have been shown to have a low-rank property: well defined off-diagonal blocks of their Schur complements can be approximated by low-rank products. Given a suitable ordering of the matrix which gives the blocks a geometrical meaning, such approximations can be computed using an SVD or a rank-revealing QR factorization. The resulting representation offers a substantial reduction of the memory requirement and gives efficient ways to perform many of the basic dense linear algebra operations.

Several strategies, mostly based on hierarchical formats, have been proposed to exploit this property. We study a simple, non-hierarchical, low-rank format called Block Low-Rank (BLR), and explain how it can be used to reduce the memory footprint and the complexity of sparse direct solvers based on the multifrontal method. We present experimental results on matrices coming from elliptic PDEs and from various other applications. We show that even if BLR based factorizations are asymptotically less efficient than hierarchical approaches, they still deliver considerable gains. The BLR format is compatible with numerical pivoting, and its simplicity and flexibility make it easy to use in the context of a general purpose, algebraic solver. This work has been published in the SIAM Journal on Scientific Computing [2].

7.14. Parallel Computation of a subset of entries of the inverse

Participants: Patrick Amestoy [INPT-IRIT, Université of Toulouse], Iain Duff [RAL and CERFACS], Jean-Yves L'Excellent, François-Henry Rouet.

We consider the computation in parallel of several entries of the inverse of a large sparse matrix. We assume that the matrix has already been factorized by a direct method and that the factors are distributed. Entries are efficiently computed by exploiting sparsity of the right-hand sides and the solution vectors in the triangular solution phase. We demonstrate that in this setting, parallelism and computational efficiency are two contrasting objectives. We develop an efficient approach and show its efficiency on a general purpose parallel multifrontal solver. This work has been published in the SIAM Journal on Scientific Computing [3].

7.15. Efficient 3D frequency-domain seismic modeling with a parallel block low-rank (BLR) direct solver

Participants: Patrick Amestoy [INPT-IRIT, University of Toulouse], Romain Brossier [ISTerre, University of Grenoble-Alpes], Alfredo Buttari [CNRS-IRIT, University of Toulouse], Jean-Yves L'Excellent, Théo Mary [UPS-IRIT, University of Toulouse], Ludovic Métivier [ISTerre-JK-CNRS], Alain Miniussi [Geoazur-CNRS-UNSA], Stéphane Operto [Geoazur-CNRS-UNSA], Alessandra Ribodetti [Geoazur-CNRS-UNSA], Jean Virieux [ISTerre-UJF, University of Grenoble-Alpes], Clément Weisbecker [INPT-IRIT, now at LSTC].

Three-dimensional frequency-domain full waveform inversion (FWI) of fixed-spread data can be efficiently performed in the visco-acoustic approximation when seismic modeling is based on a sparse direct solver. Based on the work in [3] and its extension to a parallel environment, we studied the application of a parallel algebraic Block Low-Rank (BLR) multifrontal solver providing an approximate solution of the time-harmonic wave equation with a reduced operation count, memory demand, and volume of communication relative to the full-rank solver. We analyzed the parallel efficiency and the accuracy of the solver with a realistic FWI case [19]. The application of this parallel BLR solver to a real data case from the North Sea for full waveform inversion of ocean-bottom cable data was also presented in [18], where a multiscale frequency-domain FWI is applied by successive inversions of 11 discrete frequencies in the 3.5Hz-10Hz frequency band. The velocity model built by FWI reveals short-scale features such as channels, scrapes left by drifting icebergs, fractures and deep reflectors below the reservoir level, although the presence of gas in the overburden. The quality of the FWI results is controlled by time-domain modeling and source wavelet estimation. This work was done in the context of an on-going collaboration with the Seiscope consortium (<https://seiscope2.obs.ujf-grenoble.fr/?lang=en?>).

7.16. Approximation algorithms for bipartite matching on multicore architectures

Participants: Fanny Dufossé [DOLPHIN/Inria Lille - Nord Europe], Kamer Kaya [BMI, The Ohio State Univ., USA], Bora Uçar.

We proposed [13] two heuristics for the bipartite matching problem that are amenable to shared-memory parallelization. The first heuristic is very intriguing from a parallelization perspective. It has no significant algorithmic synchronization overhead and no conflict resolution is needed across threads. We showed that this heuristic has an approximation ratio of around 0.632 under some common conditions. The second heuristic was designed to obtain a larger matching by employing the well-known Karp-Sipser heuristic on a judiciously chosen subgraph of the original graph. We showed that the Karp-Sipser heuristic always finds a maximum cardinality matching in the chosen subgraph. Although the Karp-Sipser heuristic is hard to parallelize for general graphs, we exploited the structure of the selected subgraphs to propose a specialized implementation which demonstrates very good scalability. We proved that this second heuristic has an approximation guarantee of around 0.866 under the same conditions as in the first algorithm. We discussed parallel implementations of the proposed heuristics on a multicore architecture. Experimental results, for demonstrating speed-ups and verifying the theoretical results in practice, were also provided.

7.17. Hypergraph partitioning for multiple communication cost metrics

Participants: Mehmet Deveci [BMI, The Ohio State Univ., USA], Kamer Kaya [BMI, The Ohio State Univ., USA], Umit V. Çatalyürek [BMI, The Ohio State Univ., USA], Bora Uçar.

We investigated [12] hypergraph partitioning-based methods for efficient parallelization of communicating tasks. A good partitioning method should divide the load among the processors as evenly as possible and minimize the inter-processor communication overhead. The total communication volume is the most popular communication overhead metric which is reduced by the existing state-of-the-art hypergraph partitioners. However, other metrics such as the total number of messages, the maximum amount of data transferred by a processor, or a combination of them are equally, if not more, important. Existing hypergraph-based

solutions use a two phase approach to minimize such metrics where in each phase, they minimize a different metric, sometimes at the expense of others. We proposed a one-phase approach where all the communication cost metrics can be effectively minimized in a multi-objective setting and reductions can be achieved for all metrics together. For an accurate modeling of the maximum volume and the number of messages sent and received by a processor, we proposed the use of directed hypergraphs. The directions on hyperedges necessitate revisiting the standard partitioning heuristics. We did so and proposed a multi-objective, multi-level hypergraph partitioner. The partitioner takes various prioritized communication metrics into account, and optimizes all of them together in the same phase. Compared to the state-of-the-art methods which only minimize the total communication volume, we showed on a large number of problem instances that the new method produced better partitions in terms of several communication metrics.

7.18. Comments on the hierarchically structured bin packing problem

Participants: Thomas Lambert [Inria Bordeaux Sud-Ouest], Loris Marchal, Bora Uçar.

We studied [16] the hierarchically structured bin packing problem. In this problem, the items to be packed into bins are at the leaves of a tree. The objective of the packing is to minimize the total number of bins into which the descendants of an internal node are packed, summed over all internal nodes. We investigated an existing algorithm and made a correction to the analysis of its approximation ratio. Further results regarding the structure of an optimal solution and a strengthened inapproximability result were given.

7.19. Semi-two-dimensional partitioning for parallel sparse matrix-vector multiplication

Participants: Enver Kayaaslan, Cevdet Aykanat [Bilkent Univ., Turkey], Bora Uçar.

We proposed [31] a novel sparse matrix partitioning scheme, called semi-two-dimensional (s2D), for efficient parallelization of sparse matrix-vector multiply (SpMV) operations on distributed memory systems. In s2D, matrix nonzeros are more flexibly distributed among processors than one dimensional (rowwise or columnwise) partitioning schemes. Yet, there is a constraint which renders s2D less flexible than two-dimensional (nonzero based) partitioning schemes. The constraint is enforced to confine all communication operations in a single phase, as in 1D partition, in a parallel SpMV operation. In a positive view, s2D thus can be seen as being close to 2D partitions in terms of flexibility, and being close to 1D partitions in terms of computation/communication organization. We described two methods that take partitions on the input and output vectors of SpMV and produce s2D partitions while reducing the total communication volume. The first method obtains optimal total communication volume, while the second one heuristically reduces this quantity and takes computational load balance into account. We demonstrated that the proposed partitioning method improves the performance of parallel SpMV operations both in theory and practice with respect to 1D and 2D partitionings.

7.20. Combining backward and forward recovery to cope with silent errors in iterative solvers

Participants: Massimiliano Fasi [Univ Manchester, UK], Julien Langou [Univ. Colorado Denver, USA], Yves Robert, Bora Uçar.

We proposed combining checkpointing and verification for coping with silent errors in iterative solvers. We used algorithm based fault tolerance for error detection and error correction, allowing a forward recovery (and no rollback nor re-execution) when a single error is detected. We introduced an abstract performance model to compute the performance of all schemes, and we instantiated it using the Conjugate Gradient (CG) algorithm. Finally, we validate our new approach through a set of simulations both in normal and preconditioned CG [48], [25], [47].

7.21. Load-balanced local time stepping for large-scale wave propagation

Participants: Max Rietmann [Univ. Lugano, CH], Daniel Peter [Univ. Lugano, CH], Olaf Schenk [Univ. Lugano, CH], Bora Uçar, Marcus J. Grote [Univ. Basel, CH].

In complex acoustic or elastic media, finite element meshes often require regions of refinement to honor external or internal topography, or small-scale features. These localized smaller elements create a bottleneck for explicit time-stepping schemes due to the Courant-Friedrichs-Lewy stability condition. Recently developed local time stepping (LTS) algorithms reduce the impact of these small elements by locally adapting the time-step size to the size of the element. The recursive, multi-level nature of our LTS scheme introduces an additional challenge, as standard partitioning schemes create a strong load imbalance across processors. We examined [33] the use of multi-constraint graph and hypergraph partitioning tools to achieve effective, load-balanced parallelization. We implemented LTS-Newmark in the seismology code SPECFEM3D and compared performance and scalability between different partitioning tools on CPU and GPU clusters using examples from computational seismology.

7.22. Fast and high quality topology-aware task mapping

Participants: Mehmet Deveci [BMI, The Ohio State Univ., USA], Kamer Kaya [BMI, The Ohio State Univ., USA], Umit V. Çatalyürek [BMI, The Ohio State Univ., USA], Bora Uçar.

Considering the large number of processors and the size of the interconnection networks on exascale-capable supercomputers, mapping concurrently executable and communicating tasks of an application is a complex problem that needs to be dealt with care. For parallel applications, the communication overhead can be a significant bottleneck on scalability. Topology-aware task-mapping methods that map the tasks to the processors (i.e., cores) by exploiting the underlying network information are very effective to avoid, or at worst bend, this limitation. We proposed [24] novel, efficient, and effective task mapping algorithms employing a graph model. The experiments showed that the methods are faster than the existing approaches proposed for the same task, and on 4096 processors, the algorithms improved the communication hops and link contentions by 16% and 32%, respectively, on the average. In addition, they improved the average execution time of a parallel SpMV kernel and a communication-only application by 9% and 14%, respectively.

7.23. Distributed memory tensor computations

Participants: Oguz Kaya, Bora Uçar.

There are two prominent tensor decomposition formulations. CANDECOMP/PARAFAC (CP) formulation approximates a tensor as a sum of rank-one tensors. *Tucker* formulation approximates a tensor with a core tensor multiplied by a matrix along each mode. Both of these formulations have uses in applications. The most common algorithms for both decompositions are based on the alternating least squares method. The algorithms of this type are iterative, where the computational core of an iteration is a special operation operation between an N -mode tensor and N matrices. These key operations are called the matricized tensor times Khatri-Rao product (MTTKRP) in the CP-ALS case, and the n -mode product in the Tucker decomposition case. We have investigated efficient parallelizations of full fledged algorithms for obtaining these two decompositions in distributed memory systems [30], [51] with a special focus on the mentioned key operations. In both studies, hypergraphs are used for computational load balancing and communication cost reduction. We are currently finalizing our last touches on the Tucker decomposition algorithms [51] to submit it to a conference. We are also working towards a unified view of the parallelization of the two algorithms. This work with its whole extend is carried out in the context of the thesis of Oguz Kaya.

7.24. Bridging the gap between performance and bounds of Cholesky factorization on heterogeneous platforms

Participants: Emmanuel Agullo [Inria Bordeaux Sud-Ouest], Olivier Beaumont [Inria Bordeaux Sud-Ouest], Lionel Eyraud-Dubois [Inria Bordeaux Sud-Ouest], Julien Herrmann, Suraj Kumar [Inria Bordeaux Sud-Ouest], Loris Marchal, Samuel Thibault [Inria Bordeaux Sud-Ouest].

In this work, we consider the problem of allocating and scheduling dense linear application on fully heterogeneous platforms made of CPUs and GPUs. More specifically, we focus on the Cholesky factorization since it exhibits the main features of such problems. Indeed, the relative performance of CPU and GPU highly depends on the sub-routine: GPUs are for instance much more efficient to process regular kernels such as matrix-matrix multiplications rather than more irregular kernels such as matrix factorization. In this context, one solution consists in relying on dynamic scheduling and resource allocation mechanisms such as the ones provided by PaRSEC or StarPU. We analyze the performance of dynamic schedulers based on both actual executions and simulations, and we investigate how adding static rules based on an offline analysis of the problem to their decision process can indeed improve their performance, up to reaching some improved theoretical performance bounds which we introduce [17].

7.25. Assessing the cost of redistribution followed by a computational kernel: Complexity and performance results

Participants: Julien Herrmann, George Bosilca [University of Tennessee, Knoxville], Thomas Héroult [University of Tennessee, Knoxville], Loris Marchal, Yves Robert, Jack Dongarra [University of Tennessee, Knoxville].

The classical redistribution problem aims at optimally scheduling communications when reshuffling from an initial data distribution to a target data distribution. This target data distribution is usually chosen to optimize some objective for the algorithmic kernel under study (good computational balance or low communication volume or cost), and therefore to provide high efficiency for that kernel. However, the choice of a distribution minimizing the target objective is not unique. This leads to generalizing the redistribution problem as follows: find a re-mapping of data items onto processors such that the data redistribution cost is minimal, and the operation remains as efficient. This work studies the complexity of this generalized problem. We compute optimal solutions and evaluate, through simulations, their gain over classical redistribution. We also show the NP-hardness of the problem to find the optimal data partition and processor permutation (defined by new subsets) that minimize the cost of redistribution followed by a simple computational kernel. Finally, experimental validation of the new redistribution algorithms are conducted on a multicore cluster, for both a 1D-stencil kernel and a more compute-intensive dense linear algebra routine.

This work has been published in the *Parallel Computing* journal [15].

7.26. STS-k: A Multi-level Sparse Triangular Solution Scheme for NUMA Multicores

Participants: Humayun Kabir [Penn State University], Joshua Booth [Sandia National Laboratories], Guillaume Aupy, Anne Benoit, Yves Robert, Padma Raghavan [Penn State University].

We consider techniques to improve the performance of parallel sparse triangular solution on non-uniform memory architecture multicores by extending earlier coloring and level set schemes for single-core multiprocessors. We develop STS-k, where k represents a small number of transformations for latency reduction from increased spatial and temporal locality of data accesses. We propose a graph model of data reuse to inform the development of STS-k and to prove that computing an optimal cost schedule is NP-complete. We observe significant speed-ups with STS-3 on 32-core Intel Westmere-EX and 24-core AMD ‘MagnyCours’ processors. Incremental gains solely from the 3-level transformations in STS-3 for a fixed ordering, correspond to reductions in execution times by factors of 1.4 (Intel) and 1.5 (AMD) for level sets and 2 (Intel) and 2.2 (AMD) for coloring. On average, execution times are reduced by a factor of 6 (Intel) and 4 (AMD) for STS-3 with coloring compared to a reference implementation using level sets.

This work has been published in SC’15 [29].

7.27. Mono-parametric Tiling

Participants: Guillaume Iooss [Inria/ENS-Lyon/UCBL/CNRS], Sanjay Rajopadhye [Colorado State University], Christophe Alias, Yun Zou [Colorado State University].

Tiling is a crucial program transformation with many benefits: it improves locality, exposes parallelism, allows for adjusting the ops-to-bytes balance of codes, and can be applied at multiple levels. Allowing tile sizes to be symbolic parameters at compile time has many benefits, including efficient autotuning, and run-time adaptability to system variations. For polyhedral programs, parametric tiling in its full generality is known to be non-linear, breaking the mathematical closure properties of the polyhedral model. Most compilation tools therefore either avoid it by only performing fixed size tiling, or apply it in only the final, code generation step. Both strategies have limitations.

We first introduce mono-parametric partitioning, a restricted parametric, tiling-like transformation which can be used to express a tiling. We show that, despite being parametric, it is a polyhedral transformation. We first prove that applying mono-parametric partitioning (i) to a polyhedron yields a union of polyhedra, and (ii) to an affine function produces a piecewise-affine function. We then use these properties to show how to partition an entire polyhedral program, including one with reductions. Next, we generalize this transformation to tiles with arbitrary tile shapes that can tessellate the iteration space (e.g., hexagonal, trapezoidal, etc). We show how mono-parametric tiling can be applied at multiple levels, and enables a wide range of polyhedral analyses and transformations to be applied.

This work has been published as an Inria research report [49] and will be submitted to a journal.

7.28. Data-aware Process Networks

Participants: Christophe Alias, Alexandru Plesco [XtremLogic SAS].

High-level circuit synthesis (HLS, high-level synthesis) consists in compiling a C-like high-level program to a circuit. The circuit must be as efficient as possible while using properly the resources (energy, memory, FPGA building blocks, etc). Though many progresses were achieved on the low aspects of circuit generation (pipeline, place/route), the front-end aspects (parallelism, communications) are still rudimentary compared to the state-of-the-art techniques in the HPC community.

We introduce the Data-aware Process Networks (DPN), a new parallel execution model adapted to the hardware constraints of high-level synthesis, where the data transfers are made explicit. We show that the DPN model is consistent in the meaning where any translation of a sequential program produces an equivalent DPN without deadlocks. Finally, we show how to compile a sequential program to a DPN and how to optimize the input/output and the parallelism.

This work was published as an Inria research report [63] and will be submitted to a journal.

7.29. Termination of C programs

Participants: Laure Gonnord, David Monniaux [CNRS/VERIMAG], Gabriel Radanne [Univ Paris 7/ PPS].

We designed a complete method for synthesizing lexicographic linear ranking functions (and thus proving termination), supported by inductive invariants, in the case where the transition relation of the program includes disjunctions and existentials (large block encoding of control flow).

Previous work would either synthesize a ranking function at every basic block head, not just loop headers, which reduces the scope of programs that may be proved to be terminating, or expand large block transitions including tests into (exponentially many) elementary transitions, prior to computing the ranking function, resulting in a very large global constraint system. In contrast, our algorithm incrementally refines a global linear constraint system according to extremal counterexamples: only constraints that exclude spurious solutions are included.

Experiments with our tool Termite 6.5 show marked performance and scalability improvements compared to other systems.

This work has been published in the proceedings of PLDI'15 [27].

7.30. Analysing C programs with arrays

Participants: Laure Gonnord, David Monniaux [CNRS/VERIMAG].

Automatically verifying safety properties of programs is hard, and it is even harder if the program acts upon arrays or other forms of maps. Many approaches exist for verifying programs operating upon Boolean and integer values (e.g. abstract interpretation, counterexample-guided abstraction refinement using interpolants), but transposing them to array properties has been fraught with difficulties.

In contrast to most preceding approaches, we do not introduce a new abstract domain or a new interpolation procedure for arrays. Instead, we generate an abstraction as a scalar problem and feed it to a preexisting solver. The intuition is that if there is a proof of safety of the program, it is likely that it can be expressed by elementary steps between properties involving only a small (tunable) number N of cells from the array.

Our transformed problem is expressed using Horn clauses over scalar variables, a common format with clear and unambiguous logical semantics, for which there exist several solvers. In contrast, solvers directly operating over Horn clauses with arrays are still very immature.

An important characteristic of our encoding is that it creates a nonlinear Horn problem, with tree unfoldings, contrary to the linear problems obtained by flatly encoding the control-graph structure. Our encoding thus cannot be expressed by encoding into another control-flow graph problem, and truly leverages the Horn clause format.

Experiments with our prototype VAPHOR show that this approach can prove automatically the functional correctness of several classical examples of the literature, including *selection sort*, *bubble sort*, *insertion sort*, as well as examples from previous articles on array analysis.

This work has been published as a research report [53] and is currently under submission.

7.31. Symbolic Range Analysis of Pointers in C programs

Participants: Maroua Maalej, Vitor Paisante [Univ. Mineas Gerais, Brasil], Laure Gonnord, Fernando Pereira [Univ. Mineas Gerais, Brasil], Vitor Paisante [Univ. Mineas Gerais, Brasil].

Alias analysis is one of the most fundamental techniques that compilers use to optimize languages with pointers. However, in spite of all the attention that this topic has received, the current state-of-the-art approaches inside compilers still face challenges regarding precision and speed. In particular, pointer arithmetic, a key feature in C and C++, is yet to be handled satisfactorily. We designed a new alias analysis algorithm to solve this problem. The key insight of our approach is to combine alias analysis with symbolic range analysis. This combination lets us disambiguate fields within arrays and structs, effectively achieving more precision than traditional algorithms. To validate our technique, we have implemented it on top of the LLVM compiler. Tests on a vast suite of benchmarks show that we can disambiguate several kinds of C idioms that current state-of-the-art analyses cannot deal with. In particular, we can disambiguate 1.35x more queries than the alias analysis currently available in LLVM. Furthermore, our analysis is very fast: we can go over one million assembly instructions in 10 seconds.

This work has been published at CGO'16 [32].

An extended version of the related work has also been published as an Inria research report [52] and will be the basis of a journal submission.

SCALE Team

7. New Results

7.1. Programming Languages for Distributed Systems

7.1.1. Multi-active Objects

Participants: Ludovic Henrio, Justine Rochas, Vincenzo Mastandrea.

The active object programming model is particularly adapted to easily program distributed objects: it separates objects into several *activities*, each manipulated by a single thread, preventing data races. However, this programming model has its limitations in terms of expressiveness – risk of deadlocks – and of efficiency on multicore machines. We proposed to extend active objects with *local multi-threading*. We rely on declarative *annotations* for expressing potential concurrency between requests, allowing easy and high-level expression of concurrency. This year we realized the following:

- We proved the correctness of our compiler from ABS language into ProActive multi-active objects. This translation can be generalised to many other active object languages. This work has been published as a research report, and is under submission to a conference. The proof brought us very deep and interesting understanding on the differences between the languages.
- We started to work on static detection of deadlocks for multi-active object. This is the work of Vincenzo Mastandrea who is starting a Labex PhD in collaboration with the FOCUS EPI (Univ of Bologna). An article is currently submitted to a conference on this subject.
- We are formalising in Isabelle/HOL a first version of the semantics of multiactive objects. This work was done in collaboration with Florian Kammuller
- We organised a workshop on active object languages with the main teams in Europe involved in the development of active-object languages. A journal survey paper on the subject is currently being written.
- We implemented a debugger for multi active object programs.

We plan to continue to improve the model, especially about compile-time checking of annotations and about fault tolerance of multiactive objects.

7.1.2. Behavioural Semantics

Participants: Ludovic Henrio, Eric Madelaine, Min Zhang, Siqi Li.

We are conducting a large study on Parameterised Networks of Automata (pNets) from a theoretical perspective. We started last year with some 'pragmatic' expressiveness of the pNets formalism, showing how to express a wide range of classical constructs of (value-passing) process calculi, but also complex interaction patterns used in modern distributed systems. After publishing those results [13], we focused on open systems and our formalism is able to represent operators of composition of processes, they are represented as hierarchically composed automata with holes and parameters. We defined a semantics for open pNets and a bisimulation theory for them. This study was driven by several usecase examples including a hierarchical broadcast algorithm and several operators of concurrent processes. A short presentation is accepted for publication in the journal "Science China: Information Sciences -". A full paper on the subject of the semantics and bisimulation for open pNets is under submission to a conference.

In parallel we have started the study of a denotational semantics for open pNets, based on the Universal Theory of Processes (UTP). The idea in the long term would be to draw links between the operational, denotational, and algebraic models of the pNet formalism. A short presentation of our preliminary results will be presented at the conference PDP'16 (work in progress session).

7.1.3. GPU-based High Performance Computing for finance

Participants: Michael Benguigui, Françoise Baude.

We have pursued our work on pricing American multi-dimensional (so very computation intensive) options in finance and we have been able to extend this to the computation of Value At Risk (consists in repeating the American option pricing, but we have found a financial grounded optimization that avoids us to replicate the most time consuming phase).

Moreover, the balancing of work is taking in consideration the heterogeneous nature of the involved GPUs, and is capable to harness the computing power of multi-core CPUs that also support running OpenCL codes. As our scheduling solution is capable to get a reasonable prediction of the workload of each slave computation, we have leveraged this to run the whole pricing and VaR computations on hybrid and heterogeneous clusters. These last results have been incorporated in the PhD thesis of M. Benguigui.

7.1.4. Scalable and robust Middleware for distributed event based computing

Participants: Maeva Antoine, Fabrice Huet, Françoise Baude.

In the context of the FP7 STREP PLAY and French SocEDA ANR research projects terminated late 2013, we initiated and pursued the design and development of the Event Cloud.

As a distributed system handling huge amount of information, this middleware can suffer from data imbalances. In a journal extension of a previous workshop paper [6], we have enlarged our literature review of structured peer to peer systems regarding the way they handle load imbalance to the case of distributed big data systems. We have generalized those popular approaches by proposing a core API that we have proved to be indeed also applicable to the Event Cloud middleware way of implementing a load balancing policy.

7.1.5. Vercors: Integrated environment for verifying and running distributed components

Participants: Ludovic Henrio, Oleksandra Kulankhina, Eric Madelaine.

It is the general purpose of the Vercors platform to target the generation of distributed applications with safety guarantees. In Vercors, the approach starts from graphical specification formalisms allowing the architectural and behavioral description of component systems. From this point, the user can automatically verify application properties using model-checking techniques. Finally, the specified and verified component model can be translated into executable Java code. The Vercors tool suite is distributed as an Eclipse plugin. This year

- we implemented a first reliable version of the whole tool chain including generation of verifiable models and executable Java code.
- We applied the approach to several examples including Peterson's leader election algorithm, a workflow executor, and the control and management of service composition [7].
- A paper accepted at FASE'2016 presents an overview of this work; a research report provides the full version of the paper [20]. The theoretical background was published as a research report and an improve version is being submitted as a journal paper.

The practical implementation allowed us to improve the presentation of the theory and better evaluate it.

7.2. Run-time/middle-ware level

7.2.1. Virtual Machines Scheduling

Participants: Fabien Hermenier, Vincent Kherbache.

In [19], we present BtrPlace as an application of the dynamic bin packing problem with a focus on its dynamic and heterogeneous nature. We advocate flexibility to answer these issues and present the theoretical aspects of BtrPlace and its modeling using Constraint Programming.

We also continued our work on scheduling VM migrations. In [14], [17], we propose a model for VM migration that consider their memory workload and the network topology. This model was then implemented in place of the previous migration scheduler in BtrPlace. Experiments on a real testbed show the new scheduler outperforms state-of-the-art approaches that cap the migration parallelism by a constant to reduce the completion time. Besides an optimal capping, it reduces the migration duration by 20.4% on average and the completion time by 28.1%. In a maintenance operation involving 96 VMs to migrate between 72 servers, it saves 21.5% Joules against the native BtrPlace. Finally, its current library of 6 constraints allows administrators to address temporal and energy concerns, for example to adapt the schedule and fit a power budget.

Finally, in [10] we transfer the principles of using Constraint Programming to propose a multi-objective job placement algorithm devoted to High Performance Computing (HPC). One of the key decisions made by both MapReduce and HPC cluster management frameworks is the placement of jobs within a cluster. To make this decision, they consider factors like resource constraints within a node or the proximity of data to a process. However, they fail to account for the degree of collocation on the cluster's nodes. A tight process placement can create contention for the intra-node shared resources, such as shared caches, memory, disk, or network bandwidth. A loose placement would create less contention, but exacerbate network delays and increase cluster-wide power consumption. Finding the best job placement is challenging, because among many possible placements, we need to find one that gives us an acceptable trade-off between performance and power consumption. We then propose to tackle the problem via multi-objective optimization. Our solution is able to balance conflicting objectives specified by the user and efficiently find a suitable job placement.

7.3. Application level

7.3.1. DEVS-based Modeling & Simulation

Participants: Olivier Dalle, Damian Vicino.

DEVS is a formalism for the specification of discrete-event simulation models, proposed by Zeigler in the 70's, that is still the subject of many research in the simulation community. Surprisingly, the problem of representing the time in this formalism has always been somehow neglected, and most DEVS simulators keep using Floating Point numbers for their arithmetics on time values, which leads to a range of systematic errors, including severe ones such as breaking the causal relations in the model.

In [15] we propose simulation algorithms, based on the Discrete Event System Specification (DEVS) formalism, that can be used to simulate and obtain every possible output and state trajectories of simulations that receive input values with uncertainty quantification. Then, we present a subclass of DEVS models, called Finite Forkable DEVS (FF-DEVS), that can be simulated by the proposed algorithms. This subclass ensures that the simulation is forking only a finite number of processes for each simulation step. Finally, we discuss the simulation of a traffic light model and show the trajectories obtained when it is subject to input uncertainty.

We have also worked on improving the simulation of DEVS models in some particular situations[16]. Parallel Discrete Event System Specification (PDEVs), for example, is a well-known formalism used to model and simulate Discrete Event Systems. This formalism uses an abstract simulator that defines a set of abstract algorithms that are parallel by nature. To implement simulators using these abstract algorithms, several architectures were proposed. Most of these architectures follow distributed approaches that may not be appropriate for single core processors or microcontrollers. In order to reuse efficiently PDEVs models in this type of systems, we define a new architecture that provides a single threaded execution by passing messages in a call/return fashion to simplify the execution time analysis.

This work has also been presented and defended in the PhD Thesis of D. Vicino[5].

7.3.2. Simulation of Software-Defined Networks

Participants: Olivier Dalle, Damian Vicino.

Software Defined Networks (SDN) is a new technology that has gained a lot of attention recently. It introduces programmatic ways to reorganize the network logical topology. To achieve this, the network interacts with a set of controllers, that can dynamically update the configuration of the network routing equipments based on the received events. As often with new network technologies, discrete-event simulation proves to be an invaluable tool for understanding and analyzing the performance and behavior of the new systems. In [8], we use such simulations for evaluating the impact of Software-Defined Networks' Reactive Routing on BitTorrent performance. Indeed, BitTorrent uses choking algorithms that continuously open and close connections to different peers. Software Defined Networks implementing Reactive Routing may be negatively affecting the performances of the system under specific conditions because of its lack of knowledge of BitTorrent strategies.

SOCRATE Project-Team

6. New Results

6.1. Flexible Radio Front-End

The contributions of members of this axis are mainly on four topics: Wake-Up Radio, Full-Duplex transceivers, SDR Gateways for Urban Networks, and Channel Estimation. In the global concept of enhancing wireless communications, those four topics are complimentary, addressing the reduction of energy consumption, the increase of throughput and/or flexibility of the transmission and the performance evaluation.

6.1.1. Wake-Up Radio

The last decades have been really hungry in new ways to reduce energy consumption. That is especially true when talking about wireless sensor networks in general and home multimedia networks in particular, since electrical energy consumption is the bottleneck of the network. One of the most energy-consuming functional block of an equipment is the radio front end, and methods to switch it off during the time intervals where it is not active must be implemented. This study [10] has proposed a wake-up radio circuit which is capable of both addressing and waking up not only a more efficient but also more energy-consuming radio front end. By using a frequency footprint to differentiate each sensor, awaking all the sensors except for the one of interest is avoided. The particularity of the proposed wake-up receiver is that the decision is taken in the radio-frequency part and no baseband treatment is needed. The global evaluation in theory and in simulation was performed, and a first testbed of this technology was fabricated.

6.1.1.1. Full-Duplex

This work studies [8] a Full-Duplex Dual-Band (FDDDB) OFDM radio architecture that enables the radio transceiver to be more flexible and provides a viable radio link capacity gain. A simple but practical I/Q imbalance estimation and compensation method, based on the frequency-flat-fading behavior of the self-interference channel, is proposed. The performance of the proposed I/Q imbalance compensation method is evaluated by system level simulations conducted with ADS and Matlab. The co-simulation results show that the proposed radio transceiver could potentially increase the physical layer transmission rate by four times compared to the conventional radio link at the cost of tolerable loss of BER performance. The I/Q imbalance compensation method can effectively compensate both high and low I/Q imbalance without the problem of algorithm convergence. Application of this technique for physical layer security has already been proposed.

6.1.1.2. SDR for SRDs

The technologies employed in urban sensor networks are permanently evolving, and thus the gateways of these networks have to be regularly upgraded. The existing method to do so is to stack-up receivers dedicated to one communication protocol. However, this implies to have to replace the gateway every time a new protocol is added to the network. A more practical way to do this is to perform a digitization of the full band and to perform digitally the signal processing, as done in Software-Defined Radio (SDR). The main hard point in doing this is the dynamic range of the signals: indeed the signals are emitted with very different features because of the various propagation conditions. It has been proved that the difference of power between two signals can be so important that no existing Analog-to-Digital Converter (ADC) is able to properly digitize the signals. We propose a solution to reduce the dynamic range of signals before digital conversion. In this study [28], the assumption is made that there is one strong signal, and several weak signals. This assumption is made from the existing urban sensor networks topology. A receiver architecture with two branches is proposed with a "Coarse Digitization Path" (CDP) and a "Fine Digitization Path" (FDP). The CDP allows to digitize the strong signal and to get data on it that is used to reconfigure the FDP. The FDP then uses a notch filter to attenuate the strong signal (and then to reduce the dynamic range of the signals) and digitizes the rest of the band. Another way to relax these specifications on ADCs is an analog processing, such as companding, that should be performed before digitization. The companding technique is usually employed on one signal (and

not on multiple signals that are only separated on the frequency domain). This work [36], [29] studies three companding laws to test their efficiency in relaxing the digitization constraints with multiple signals. A μ -law, a Piecewise-Linear (PL) law and a Piecewise-Linear, Constant Gain with Offsets (PLCGO) law are tested. We have described how to use a PLCGO approach to reduce ADC's complexity, and two implementations of the compressing law are proposed.

6.1.1.3. Channel Estimation

In modern mobile telecommunications, shadow fading has to be modeled by a two-dimensional (2D) correlated random variable since shadow fading may present both cross-correlation and spatial correlation due to the presence of similar obstacles during the propagation. In our study, 2D correlated random shadowing is generated based on the multi-resolution frequency domain ParFlow (MR-FDPF) model. The MR-FDPF model is a 2D deterministic radio propagation model, so a 2D deterministic shadowing can be firstly extracted from it. Then, a 2D correlated random shadowing can be generated by considering the extracted 2D deterministic shadowing to be a realization of it. Moreover, based on the generated 2D correlated random shadowing, a complete 2D semi-deterministic path loss model can be proposed. The proposed methodology [5] can be implemented into system-level simulators where it will be very useful due to its ability to generate realistic shadow fading.

[23] presents the first implementation on software defined radio nodes in the large scale testbed CorteXlab of a radio link estimation technique based on OFDM transmissions. The purpose of this large scale testbed is to offer to the whole scientific community an open tool to test new techniques for multiuser, cooperative and cognitive radio networks in a controlled environment. As the experimentation room was defined in order to offer reproducible measurements, it is important to be able to characterize each radio link between all transceivers. Therefore, we present here the development of a channel sounder directly implemented on the software radio nodes. This paper presents the first implementation on software defined radio nodes in the large scale testbed called CorteXlab of a radio link estimation technique based on OFDM transmissions. The purpose of this large scale testbed is to offer to the whole scientific community an open tool to test new techniques for multiuser, cooperative and cognitive radio networks in a controlled environment. As the experimentation room was defined in order to offer reproducible measurements, it is important to be able to characterize each radio link between all transceivers. Therefore, we proposed the development of a channel sounder directly implemented on the software radio nodes.

6.2. Agile Radio Resource Sharing

This axis addresses the challenges relative to the network perspective of software radio. While the two other axes have their focus on the design of the software radio nodes, axis 2 deals with coexistence and cooperation in a multi-user communications perspective.

A first research direction concerns theoretical limits of different reference scenarios where trade-offs between spectral efficiency, energy efficiency, stability and/or fairness are analyzed. This work exploits multi-users information theory, game theory and stochastic geometry. This year, a particular focus has been put on the interference channel with feedback and on dense wireless networks. New problems have been also investigated with the simultaneous energy and information transmission problem for energy harvesting and some specific attacks in smart grids.

In parallel our research activities are also driven by applicative frameworks. Concerning 4G RAN, a new interference alignment scheme has been proposed, simulated and implemented on CortexLab. This work has been presented as one of the promising technologies proposed by Greentouch. IoT has been identified as a new challenge for 5G with the objective of serving a very large number of nodes per cell, in a connectionless manner and with very small packets. The original transmission technology using ultra narrow band modulation and proposed by Sigfox for large area of IoT nodes has been investigated. A multiband CSMA strategy has been also evaluated in collaboration with CEA-Leti for dense Wifi like IoT access networks. Body area networks (BANs) represent also a very challenging applicative framework, with strong dynamics, interference environments, and low energy requirements. In partnership with Euromedia and Hikob, our studies focused on

dynamic algorithms for information gathering in a sport event broadcast system. Additionally, localization capabilities at the body scale may offer interesting perspectives but require specific MAC protocols.

6.2.1. Fundamental Limits

6.2.1.1. Energy efficiency - Spectral Efficiency (EE-SE) Tradeoffs in Wireless RANs

The spectral and energy efficiency (SE-EE) trade-off in cellular networks has attracted significant recent interest in the wireless community [1]. The work in [7] studies this fundamental limit with a simple and effective method. The proposed theoretical framework is based on an optimal radio resource allocation of transmit power and bandwidth for the downlink direction, applicable for an orthogonal cellular network. The analysis is initially focused on a single cell scenario, for which in addition to the solution of the main SE-EE optimization problem, it is proved that a traffic repartition scheme can also be adopted as a way to simplify this approach. By exploiting this interesting result along with properties of stochastic geometry, this work is extended to a more challenging multi-cell environment, where interference is shown to play an essential role and for this reason several interference reduction techniques are investigated. Special attention is also given to the case of low signal to noise ratio (SNR) and a way to evaluate the upper bound of EE in this regime is provided. This methodology leads to tractable analytical results under certain common channel properties, and thus allows the study of various models without the need for demanding system level simulations.

6.2.1.2. Interference Channels with Feedback

The capacity region of the two-user linear deterministic (LD) interference channel with noisy output feedback (IC-NOF) is fully characterized in [35], [26]. This result allows the identification of several asymmetric scenarios in which implementing channel-output feedback in only one of the transmitter-receiver pairs is as beneficial as implementing it in both links, in terms of achievable individual rate and sum-rate improvements w.r.t. the case without feedback. In other scenarios, the use of channel-output feedback in any of the transmitter-receiver pairs benefits only one of the two pairs in terms of achievable individual rate improvements or simply, it turns out to be useless, i.e., the capacity regions with and without feedback turn out to be identical even in the full absence of noise in the feedback links. As a byproduct, the exact conditions on the signal to noise ratios on the feedback links to observe an improvement on either a single rate, both single rates, or the sum-rate capacity, for any IC-NOF are also fully described in [41].

6.2.1.3. Simultaneous Energy and Information Transmission

The fundamental limits of simultaneous information and energy transmission in the two-user Gaussian multiple access channel (G-MAC) with and without feedback are fully characterized in [33], [9]. All the achievable information and energy transmission rates (in bits per channel use and energy-units per channel use respectively) are identified. Thus, the information-energy capacity region is defined in both cases. In the case without feedback, an achievability scheme based on power-splitting and successive interference cancellation is shown to be optimal. Alternatively, in the case with feedback (G-MAC-F), a simple yet optimal achievability scheme based on power-splitting and Ozarow's capacity achieving scheme is presented. Three of the most important observations in this work are: (a) The capacity-energy region of the G-MAC without feedback is a proper subset of the capacity-energy region of the G-MAC-F; (b) Feedback can at most double the energy rate for a fixed information rate; and (c) Time-sharing with power control is strictly suboptimal in terms of sum-rate in the G-MAC without feedback.

6.2.1.4. Multiple Access Channel and Broadcast Channel with Linear Feedback Schemes

In [11], it is shown that for the two-user Gaussian broadcast channel with correlated noises and perfect feedback the largest region that can be achieved by linear-feedback schemes equals the largest region that can be achieved over a dual multi-access channel when in this latter the channel inputs are subject to a "non-standard" sum-power constraint that depends on the BC-noise correlation. Combining this new duality result with Ozarow's MAC-scheme gives an elegant achievable region for the Gaussian BC with correlated noises. A constructive iterative coding scheme is then presented for the non-symmetric Gaussian BC with uncorrelated noises that is sum-rate optimal among all linear-feedback schemes. This coding scheme shows that the connection between the MAC and the BC optimal schemes is tighter than what is suggested by our duality result on achievable rates. In fact, it is linear-feedback sum-rate optimal to use Ozarow MAC-encoders and MAC-decoders—rearranged—to code over the BC.

6.2.2. Low Complexity Receivers for Massive MIMO Systems

In wireless communications, Multi-user massive MIMO network is a scenario that has been recently proposed, where many mobile terminals are served by a Base Station (BS) equipped with a very high number of antennas. In such a scenario, the detection in the uplink remains a challenge, since the BS is required to detect signals transmitted from all users while trying to exploit full received diversity. The optimal detection criterion that fulfills the diversity requirement is the Maximum-Likelihood (ML) joint detection which has been proposed to detect jointly the transmitted signals. However, such a criterion is not applicable to the addressed multi-user massive MIMO scenario due to its computational complexity that increases exponentially with the number of signals to be detected. In our work paper, we have proposed a relaxed ML detector based on an iterative decoding strategy that reduces the computational cost. We exploit the fact that the transmit constellation is discrete, and remodel the channel as a MIMO channel with sparse input belonging to the binary $\{0, 1\}$. The sparsity property allows us to relax the ML problem as a quadratic minimization under linear and ℓ_1 -norm constraint. We then prove the equivalence of the relaxed problem to a convex optimization problem solvable in polynomial time. Simulation results illustrate the efficiency of the low-complexity proposed detector compared to other existing ones in very large and massive MIMO context.

6.2.3. Distributed Radio Resource Management

6.2.3.1. Interference Alignment in Cellular Networks with no-Explicit Coordination

Current networks aim to support high data rates for end users by increasing the spectral efficiency in bits-per-Hertz, at the expense of the energy efficiency of the network. Indeed, an important part of the energy consumption of mobile networks is proportional to the radiated energy, which relies on the frequency bandwidth and the transmission power. Any energy efficient transmission scheme should exploits the whole system bandwidth by allocating the entire available spectrum to each base station. Such an approach, however, leads to significant interference increase and performance degradation for mobiles located at the cell edges. The key challenge is to balance interference avoidance and spectrum use to reach an optimal spectral efficiency – energy efficiency (EE-SE) trade-off. The work achieved in the framework of Greentouch collaboration is based on the non classical interference alignment scheme proposed by Suh and Tse in downlink mode. The key contribution relies on users scheduling with a unique criteria based as well on QoS priorities and orthogonality of precoding directions. The spectral efficiency is improved by a factor 2 for edge users and a energy saving of about 30% is made possible. This scheme has been evaluated on simulation scenarios as defined by Greentouch partners and a simplified version has been implemented on FIT/CorteXlab and demonstrastrated during the final event of Greentouch (New-York, June 2015).

6.2.4. RANs for IoT : Dense and Connectionless Solutions

Internet of Things (IoT) is going to take a major place in the telecommunications market as announced in technical and public medias. The paradigm of IoT relies on the deployment of billions of objects having the capability of transmitting information about their context and environment and to create a real-time, secured and efficient interaction between the real and the virtual worlds, pushing them to evolve from the state of cousins to the state of Siamese twins. IoT revealed to be a key technology for solving societal issues such as digital cities, intelligent transportation, green environment monitoring or medical care and elderly person monitoring.

IoT has strong connections with machine-to-machine (M2M), and sometimes in literature, both terms refer to the same idea. From our point of view, IoT covers a broader scope including as well the technology and the applications. On the opposite, M2M refers to the technologies that allow machines or objects to communicate.

In any case, from the technical point of view, the main challenge of this new paradigm is to let a huge number of machine type devices (MTDs) be connected to the Internet at a low cost, with a limited infrastructure and featuring a very long life time with very small battery or energy needs [4].

In this global picture, we may consider different technical issues. M2M has first been defined to connect MTDs in their vicinity. The proposed solutions extensively rely on the research results produced over the last 20 years for ad-hoc and wireless sensor networks. Initiated 20 years ago from theoretical concepts, this very active research area has gone up to the definition of full standards (802.15.4, 802.15.6, Zigbee, Bluetooth) which have already found a market.

More recently, the IoT paradigm has been extended to the problem of connecting all these MTDs to the Internet, and through Internet to anyone or anything. The massive connection of objects spread over the world is a challenge that has some similarities with the paradigm of cellular networks which aimed at connecting people. This similarity attracted the interest of mobile network providers, to exploit such attractive potential market and IoT has been identified as a target for the future 5G.

6.2.4.1. Performance of Ultra-NarrowBand Techniques

The Ultra-narrow-band technology is an appealing solution for the low throughput wireless sensor networks (10b/s - 1 kb/s). It is complementary to the classical cellular networks thanks to its low energy consumption and very long range communication (up to 50 km in free-space) [4]. This technology has already been deployed and is proved to be ultra-efficient for point-to-point communications in Sigfox's network. Nodes are transmitting at a random time and random frequency carrier (random frequency division multiple access schemes : R-FDMA), so the uplink is exposed to interference. In our approach, we have proposed to model this interference for the UNB network when taking into account the path-loss and Rayleigh effects, with stochastic geometry tools. The obtained model allows us to estimate the system performance, and its capacity in terms of maximum number average of simultaneous nodes in a unique cell [37]. We have also considered the replication mechanism, and identified the optimum number of replications.

6.2.4.2. Multiband CSMA for Dense Wireless Networks in Uplink

In this approach, the objective is to mitigate the degradation of the throughput and delay performance in wireless local area networks (WLAN) that employ carrier sense multiple access collision avoidance (CSMA/CA) protocol with request to send and clear to send (RTS/CTS) mechanism, when a large number of IoT like nodes are deployed. In our approach, the overhead is reduced with a modified handshake mechanism. The medium access control (MAC) overhead caused by the RTS and CTS messages is high comparing to the total duration of successful transmission. In order to reduce the MAC overhead we propose in this work a new strategy to serve many users successively. This strategy consists on sending many RTS in parallel by different stations on different frequency sub-bands. Once the RTS messages do not collide with each other, there will be no need to resend the RTS and wait for a CTS to gain the channel access [21].

6.2.5. Algorithms and Protocols for BANs

6.2.5.1. Information Gathering in a Group of Mobile Users

Distributed decisions within any group of agents, is a very active research area and theoretical results as well as efficient algorithms have already been proposed but in the context of wireless networks, the task is made harder due to possible transmission errors, channel asymmetry, dynamic behaviour of the channel and node mobility. In this work, we consider a group of mobile agents moving roughly in a common direction. We study different algorithmic solutions allowing each agent to periodically discover its neighbours: one-hop neighbours as well as multi-hop neighbors. The reference scenario is a bike race, during which groups are susceptible to split or merge. The objective is a live gathering of information about who is present in a group for live TV broadcasting. For that, we need a fully distributed approach allowing every agent to discover with a consensus algorithm the list of neighbours participating to the same pack. This study may be of interest for various other applications such as group navigation support in crowded environments, autonomous navigation of a fleet of robots. . . This problem exhibits some similarities with a clustering problem. However, a clustering problem aims at exploiting the structure of a graph and to form some subgroups to ensure a good structure of the network for further communications while our objective is rather to estimate the groups naturally formed in the real world. Hence, we have focused on distributed decision algorithms, which are widely present in the literature. Max-consensus problem has been much less studied than average consensus. The proposed algorithms are based on the N-dimension generalization of the Random Broadcast Max-Consensus algorithm,

allowing each agent to build and share the list of its multi-hop neighbors. We extend this approach to a dynamic context where the group information needs to be updated according to possible group merge or split. Experimental validation has been done in the context of a cycling race with 10 agents, equipping each bicycle with a wireless sensor node to assess the interactions between the racers and to provide a live monitoring of the dynamic evolution of the cyclists groups that form during the race.

6.2.5.2. *MAC Protocols and Algorithms for Localization at the Body Scale*

The purpose of this work is to evaluate the impact of the node speed on the ranging estimation for location applications with Wireless Body Area Networks (WBAN). While estimated with the 3-Way ranging protocol (3-WR), this distance between two nodes placed on the body can be affected by the human movements [30], [17]. Thus, we study theoretically the ranging error with the 3-WR, based on a perfect channel, a MAC layer based on TDMA using two scheduling strategies (Single node localization (P2P-B) and Aggregated & Broadcast (A&B)) and a PHY layer based on Ultra Wideband (IR-UWB) [31]. We demonstrate the accuracy of the model, and show that the distance error is highly correlated with the speed of nodes [16], while the associated mobility model has an impact on the design of MAC strategies by simulation [18].

6.2.6. *Other Topics*

6.2.6.1. *Data Injection Attacks in Smart Grids*

Multiple attacker data injection attack construction in electricity grids with minimum-mean-square-error (MMSE) state estimation is studied for centralized and decentralized scenarios in [34]. A performance analysis of the trade-off between the maximum distortion that an attack can introduce and the probability of the attack being detected by the network operator is considered. Within this setting, optimal centralized attack construction strategies are studied. The decentralized case is examined in a game-theoretic setting. A novel utility function is proposed to model this trade-off and it is shown that the resulting game is a potential game. The existence and cardinality of the corresponding set of Nash Equilibria (NE) in the game is analyzed. For the particular case of two attackers, numerical results based on IEEE test systems are presented. These results suggest that attackers perform better when they seize control of power flow measurements instead of power injection measurements.

6.3. *Software Radio Programming Model*

6.3.1. *Data Flow Programming*

Streaming languages have been proven to be a natural and efficient approach for taking advantage of the intrinsic parallelism of modern CPU architectures. The focus of many previous work has been to improve the throughput of streaming programs. In [27], we rather focus on satisfying quality-of-service requirements of streaming applications executed alongside non-streaming processes. We monitor synchronous dataflow (SDF) programs at runtime both at the application and system levels, in order to identify violations of quality-of-service requirements. Our monitoring requires the programmer to provide the expected throughput of its application (e.g 25 frames per second for a video decoder), then takes full benefit from the compilation of the SDF graph to detect bottlenecks in this graph and identify causes among processor or memory overloading. It can then be used to perform dynamic adaptations of the applications in order to optimize the use of computing and memory resources.

6.3.2. *Smart Sensors*

The article [19] presents the development of a wireless wearable sensor for the continuous, long-term monitoring of cardiac activity. Heart rate assessment, as well as heart rate variability parameters are computed in real time directly on the sensor, thus only a few parameters are sent via wireless communication for power saving. Hardware and software methods for heart beat detection and variability calculation are described and preliminary tests for the evaluation of the sensor are presented. With an autonomy of 48 hours of active measurement and a Bluetooth Low Energy radio technology, this sensor will form a part of a wireless body network for the remote mobile monitoring of vital signals in clinical applications requiring automated collection of health data from multiple patients.

6.3.3. Cryptography

For security applications in wireless sensor networks (WSNs), choosing best algorithms in terms of energy-efficiency and of small memory requirements is a real challenge because the sensor networks are composed of low-power entities. Previous works benchmarked 12 block-ciphers on an ATMEL AVR ATtiny45 8-bit microcontroller. In [2], most of the recent lightweight block cipher proposals, as well as some conventional block ciphers, are studied on the Texas Instruments MSP430 16-bit microcontroller. The chosen block ciphers are described with a security and an implementation summary. Implementations are then evaluated on a dedicated platform.

6.3.4. Hardware Arithmetic

6.3.4.1. Hardware Implementations of Fixed-Point Atan2

The atan2 function computes the polar angle $\arctan(x/y)$ of a point given by its cartesian coordinates. It is widely used in digital signal processing to recover the phase of a signal. The article [14] studies for this context the implementation of atan2 with fixed-point inputs and outputs. It compares the prevalent CORDIC shift-and-add algorithm to two multiplier-based techniques. The first one reduces the bivariate atan2 function to two functions of one variable: the reciprocal, and the arctangent. These two functions may be tabulated, or evaluated using bipartite or polynomial approximation methods. The second technique directly uses piecewise bivariate polynomial approximations, in degree 1 and degree 2. It requires larger tables but has the shortest latency. Each of these approaches requires a relevant argument reduction, which is also discussed. All the algorithms are described with the same accuracy target (faithful rounding) and implemented with similar care in FloPoCo. Based on synthesis results on FPGAs, their relevance domains are discussed.

6.3.4.2. Fixed-Point Implementations of the Reciprocal, Square Root and Reciprocal Square Root Functions

Implementations of the reciprocal, square root and reciprocal square root often share a common structure. The article [39] is a survey and comparison of methods for computing these functions. It compares classical methods (direct tabulation, multipartite tables, piecewise polynomials, Taylor-based polynomials, Newton-Raphson iterations). It also studies methods that are novel in this context: the Halley method and, more generally, the Householder method. The comparisons are made in the context of the same accuracy target (faithful rounding) and of an arbitrary fixed-point format for the inputs and outputs (precisions of up to 32 bits). Some of the methods discussed might require some form of range reduction, depending on the input range. The objective of the article is to optimize the use of fixed-size FPGA resources (block multipliers and block RAMs). The discussions and conclusions are based on synthesis results for FPGAs.

6.3.4.3. Fixed-Point Hardware Polynomials

Polynomial approximation is a general technique for the evaluation of numerical functions of one variable such as atan, reciprocal and square roots studied above. The article [38] addresses the automatic construction of fixed-point hardware polynomial evaluators. By systematically trying to balance the accuracy of all the steps that lead to an architecture, it simplifies and improves the previous body of work covering polynomial approximation, polynomial evaluation, and range reduction. This work is supported by an open-source implementation in FloPoCo.

6.3.5. Software Elementary Functions

6.3.5.1. Code Generators for Mathematical Functions

A typical floating-point environment includes support for a small set of about 30 mathematical functions such as exponential, logarithms and trigonometric functions. These functions are provided by mathematical software libraries (libm), typically in IEEE754 single, double and quad precision. The article [13] suggests to replace this libm paradigm by a more general approach: the on-demand generation of numerical function code, on arbitrary domains and with arbitrary accuracies. First, such code generation opens up the libm function space available to programmers. It may capture a much wider set of functions, and may capture even standard functions on non-standard domains and accuracy/performance points. Second, writing libm code requires fine-tuned instruction selection and scheduling for performance, and sophisticated floating-point

techniques for accuracy. Automating this task through code generation improves confidence in the code while enabling better design space exploration, and therefore better time to market, even for the libm functions. This article discusses, with examples, the new challenges of this paradigm shift, and presents the current state of open-source function code generators.

6.3.5.2. *Computing Floating-Point Logarithms with Fixed-Point Operations*

Elementary functions from the mathematical library input and output floating-point numbers. However it is possible to implement them purely using integer/fixed-point arithmetic. This option was not attractive between 1985 and 2005, because mainstream processor hardware supported 64-bit floating-point, but only 32-bit integers. Besides, conversions between floating-point and integer were costly. This has changed in recent years, in particular with the generalization of native 64-bit integer support. The purpose of the article [40] is therefore to reevaluate the relevance of computing floating-point functions in fixed-point. For this, several variants of the double-precision logarithm function are implemented and evaluated. Formulating the problem as a fixed-point one is easy after the range has been (classically) reduced. Then, 64-bit integers provide slightly more accuracy than 53-bit mantissa, which helps speed up the evaluation. Finally, multi-word arithmetic, critical for accurate implementations, is much faster in fixed-point, and natively supported by recent compilers. Novel techniques of argument reduction and rounding test are introduced in this context. Thanks to all this, a purely integer implementation of the correctly rounded double-precision logarithm outperforms the previous state of the art, with the worst-case execution time reduced by a factor 5. This work also introduces variants of the logarithm that input a floating-point number and output the result in fixed-point. These are shown to be both more accurate and more efficient than the traditional floating-point functions for some applications.

SPIRALS Project-Team

7. New Results

7.1. Traceability of Concerns in Large Software Systems

In 2015, we obtained new results in the domain of the analysis of large software systems. The purpose is to be able to deal the complexity of such systems by slicing them depending on different concerns. The slicing enables to gain a view and a better understanding on how the concern evolves over time and through the different refinement layers of the software system. For that, we present a systematic approach based on model driven engineering and basic models of software components, in order to better manage software complexity and traceability of functional and non-functional requirements. We provide in particular three major contributions. First, we provide an integrated set of meta-models for describing the concerns of software requirements, software components, and traceability between the concerns and software components. By providing an abstract model, we are independent of any implementation and thus allow existing approaches relying on that model to expand their support. With the second contribution, we propose a formal support of our model to allow formal verification. We focus on temporal property verification. For this, our design model is translated into timed automata for which we can apply a timed model checker. Instead of using temporal logic, which is difficult to handle by non-experts, we use patterns of temporal properties. For each pattern, we propose timed automata that can be applied directly into a timed model checking tool. These timed automata are seen as observers or watch dogs that check the system under observation. Finally, with the last contribution, we propose a software component-based development and verification approach, called SARA, and included in V-lifecycle widely used in the railway domain. These contributions have been validated with case studies from the domain of railway control systems especially for the new European train control system ERTMS/ETCS. These results contribute to our objective on self-optimizing software systems (see Section 3.3) and are part of the PhD thesis by Marc Sango [13].

7.2. Automatic Analysis and Repair of Exception Bugs for Java Programs

In 2015, we obtained new results in the field of automated software repair, that is a new and emerging domain of software engineering. The goal of automatic repair is to increase the quality of software systems by automatizing tasks related to fixing of defects and bugs. The new results that we bring are related with the management of runtime exceptions. These results contribute to our objective on self-healing software systems (see Section 3.2) and are part of the PhD thesis by Benoit Cornu [11], defended on 26 November 2015. To improve the available information about exceptions, we have presented a characterization of the exceptions (expected or not, anticipated or not), and of their corresponding resilience mechanisms [16]. We have provided definitions about what is a bug when facing exceptions and what are the already-in-place corresponding resilience mechanisms. We have formalized two formal resilience properties: source-independence and pure-resilience as well as an algorithm to verify them. Then, we have presented two dynamic analysis techniques based on code transformation for analyzing exceptions. Casper is an approach to make bug fixing easier by providing information about the origin of null pointer dereferences. NpeFix is a system to tolerate null pointer dereferences. Both systems are empirically validated on real-world null dereference bugs from large-scale open-source projects

STORM Team

7. New Results

7.1. Memory Management and Distributed Computing with StarPU

Task-based programming models manage to abstract away much of the architecture complexity while efficiently meeting the performance challenge, even at large scale. While computation scheduling has been well studied, the dynamic management of memory resource subscription inside such run-times has however been little explored, despite the fact that the lookahead, anticipative capabilities offered by the decoupled task submission/task execution steps may sometime come with a high memory cost, especially in distributed context where buffers for receiving incoming contributions have to be accounted for. We therefore studied the cooperation between a task-based application code and a run-time system engine to control the memory subscription levels throughout the execution. We showed that the task paradigm allows to control the memory footprint of the application by throttling the task submission flow rate, striking a compromise between the performance benefits of anticipative task submission and the resulting memory consumption.

7.2. Simulation Execution with StarPU and SimGrid

The combination of StarPU and SimGrid allows to fast, accurate, and reproducible simulation the execution of task-based HPC applications.

This has proved to be very useful for theoretical research on scheduling heuristics [10]. It notably allowed to modify the simulated platform in order to easily observe and understand which parts of the platform (bandwidth, computation power) cause a bottleneck. It also allowed to remove some parts of the problem, such as the cost of data transfers, to simplify the problem and be able to deeply study scheduling solutions and compare them with optimum solutions in a simple environment before tackling the complete platform.

We have also extended the modelization of computation nodes, to take into account the PCI hierarchy of the system. This allows to get a more accurate simulation of systems which have dedicated channels between GPUs.

Last but not least, we have started to extend the StarPU+Simgrid combination to StarPU+MPI+Simgrid, to simulate the execution of HPC applications on *clusters* of heterogeneous systems. The preliminary results seem to show good accuracy. This will allow to easily study how applications scale, and study for instance how network performance have impact on it.

7.3. Scheduling heuristics for dense linear algebra

In the context of Suraj KUMAR's PhD thesis, we are studying the scheduling of the Cholesky factorization on heterogeneous systems.

We have started to introduce communication costs into the constraint programming. Since this increases resolution time a lot, we had to optimize the expression of the data transfers to simplify the resolution. We modified the StarPU runtime system to be able to inject not only a static schedule of tasks, but also a static schedule of data transfers. This allows to inject the schedule optimized by constraint programming into real executions.

We have also shown how static schedules and dynamic scheduling strategies compare on heterogeneous platforms, and notably in the context of varying task execution time can typically be a problem for static scheduling. Static schedules have actually proven to be robust against variation in execution time. We have also studied injecting static information into dynamic schedulers, which improves the resulting performance with little offline analysis.

7.4. Out-of-core support for task graphs

In the context of the Hi-BOX project, Airbus factorizes very large compressed matrices, which can not fit in the main memory, and most of the data thus have to be temporarily transferred to the disk, and loaded on-demand. We have thus extended the StarPU out-of-core support to the case of compressed matrices, and improved the eviction heuristics, so as to transfer data to the disk in advance.

7.5. Parallel Tasks within StarPU

One of the biggest challenge raised by the design of high performance task-based applications on top of heterogeneous accelerator-based machines lies in choosing the adequate granularity of tasks. Indeed, GPUs generally exhibit better performance when executing kernels featuring numerous threads whereas regular CPU cores typically reach their peak performance with fine grain tasks working on a reduced memory footprint. As a consequence, task-based applications running on such heterogeneous platforms have to adopt an intermediate granularity, chosen as a trade-off between coarse-grain and fine-grain tasks. We have tackle this granularity problem via resource aggregation : our approach consists in reducing the performance gap between accelerators and single cores by scheduling parallel tasks over cluster of CPUs. For this purpose, we have extended the concept of scheduling context, which we introduced in a previous work, so that the main runtime system only sees virtual computing resources on which it can schedule parallel tasks (e.g. BLAS kernels). The execution of tasks inside such clusters can virtually rely on any thread-based runtime system, and does not interfere with the outer scheduler. As a proof of concept we allow the interoperability of StarPU and OpenMP to co-manage task parallelism. We showed that our approach is able to outperform the magma, dplasma and chameleon state-of-the-art dense linear algebra libraries when dealing with matrices of small and medium size.

7.6. Running Compliant OpenMP Applications on top of StarPU with the Klang-Omp Compiler

Several robust runtime systems proposed recently have shown the benefits of task-based parallelism models. However, the common weakness of these supports is to tie applications to specific APIs. The OpenMP specification aims at providing a common parallel programming means for shared-memory platforms. It appears a good candidate to address this issue. We assessed the effectiveness and limits of this approach on the ScalFMM library developed by Inria HiePACS team, implementing fast multipole methods (FMM) algorithms. We showed that OpenMP dependent tasks allow for significant performance improvements over OpenMP loops and independent tasks for this application. We also demonstrated that suitable, targeted language extensions can further improve performances by a important margin in some cases.

7.7. Task-based Seismology Simulation on top of StarPU

Understanding three-dimensional seismic wave propagation in complex media is still one of the main challenges of quantitative seismology. Because of its simplicity and numerical efficiency, the finite-differences method is one of the standard techniques implemented to consider the elastodynamics equation. Additionally, this class of modelling heavily relies on parallel architectures in order to tackle large scale geometries including a detailed description of the physics. Last decade, significant efforts have been devoted towards efficient implementation of the finite-differences methods on emerging architectures. These contributions have demonstrated their efficiency leading to robust industrial applications. The growing representation of heterogeneous architectures combining general purpose multicore platforms and accelerators leads to re-design current parallel application. We thus considered the StarPU task-based runtime system in order to harness the power of heterogeneous CPU+GPU computing nodes. Preliminary results demonstrate significant speedups in comparison with the best implementation suitable for homogeneous cores.

7.8. Interfacing the MPC Parallel Framework with StarPU

CEA has developed a framework named MPC that transforms MPI+OpenMP applications into a lightweight thread-based program which can flexibly and efficiently exploit multicore architectures. StarPU, on its side, is mainly dedicated to schedule coarse grain tasks over accelerators, and is less suited to fine grain task scheduling. We have thus initiated a software interoperability effort between StarPU and MPC. The first step was to implement a new StarPU task scheduling strategy based on a NUMA-aware adaptative task granularity according to the target device (GPU or CPU). We observed significant performance gains for a Cholesky application in comparison to an eager strategy, thanks to the NUMA-aware aspect. However more work is still needed with respect to task decomposition as it implies data partitioning during the execution. We are also working on a variable granularity task programming interface in order to simplify the developer's coding effort. Finally, we develop a mechanism in StarPU to isolate some parts of the computing platform for another runtime. We used nested *scheduling contexts* to redirect some tasks to a scheduling component that StarPU may or may not control. The idea is to associate a scheduling subcontext to a runtime, for instance MPC, that would access to a dedicated set of computing resources for executing parallel kernels.

7.9. A Domain Specific Framework for Executing Stencil Kernels on Accelerated Platforms with SYCL

Stencil kernels arise in many scientific codes as the result from discretizing natural, continuous phenomena. Many research works have designed stencil frameworks to help programmer optimize stencil kernels for performance, and to target CPUs or accelerators. However, existing stencil kernels, either library-based or language-based necessitate to write distinct source codes for accelerated kernels and for the core application, or to resort to specific keywords, pragmas or language extensions. SYCL is a C++ based approach designed by the Khronos Group to program the core application as well as the application kernels with a single unified, C++ compliant source code. A SYCL application can then be linked with a CPU-only runtime library or processed by a SYCL-enabled compiler to automatically build an OpenCL accelerated application. We designed a stencil-dedicated domain specific embedded language (DSEL) which leverage SYCL together with expression template techniques to implement statically optimized stencil applications able to run on platforms equipped with OpenCL devices, while preserving the single source benefits from SYCL. Our stencil DSL has been tested using the SYCL compiler ComputeCpp from the CodePlay company on an accelerated platform, as well as with the TriSYCL library designed as a compilerless approach for CPU-only prototyping.

7.10. Combining Code Generation and Template Specialization Techniques in the P-EDGE Generic Polar Error Correction Code Framework

Error Correction Code decoding algorithms for consumer products such as *Internet of Things* (IoT) devices are usually implemented as dedicated hardware circuits. As processors are becoming increasingly powerful and energy efficient, there is now a strong desire to perform this processing in software to reduce production costs and time to market. The recently introduced family of Successive Cancellation decoders for Polar codes has been shown in several research works to efficiently leverage the ubiquitous SIMD units in modern CPUs, while offering strong potentials for a wide range of optimizations. Together with the IMS Laboratory, we designed the P-EDGE framework which combines a specialized skeleton generator and a building blocks library routines to provide a generic, extensible Polar code exploration workbench. It enables ECC code designers to easily experiments with combinations of existing and new optimizations, while delivering performance close to state-of-art decoders.

7.11. Binary Kernel Analysis, Hinting and Transformation for SIMDization

SIMD processor units have become ubiquitous. Using SIMD instructions is the key for performance for many applications. Modern compilers have made immense progress in generating efficient SIMD code. However, they still may fail or SIMDize poorly, due to conservativeness, source complexity or missing

capabilities. When SIMDization fails, programmers are left with little clues about the root causes and actions to be taken. Our proposed guided SIMDization framework builds on the assembly-code quality assessment toolkit MAQAO to analyze binaries for possible SIMDization hindrances. It proposes improvement strategies and readily quantifies their impact, using *in vivo* evaluations of suggested transformation. Thanks to our framework, the programmer gets clear directions and quantified expectations on how to improve his/her code SIMDizability. We show results of our technique on TSVC benchmark.

7.12. Dynamic Granularity Adaptation of OpenCL Kernels on Heterogeneous Multi-device Systems

On-going work as part of the PhD of P. Huchant aims to transparently execute an OpenCL kernel, and further a complete task graph, on an heterogeneous multi-device system. We propose methods to split an OpenCL kernel at compile time and adapt its granularity dynamically to ensure load balance. If the kernel is executed multiple times, we propose to determine its granularity by using a linear program whose constraints are built from performance measurements collected during the first invocations of the kernel with predefined granularities. Splitting the execution of one kernel into different executions does not require additional information from the user, therefore increasing the level of portability of OpenCL codes. First experiments show the interest of our approach.

TACOMA Team

6. New Results

6.1. Self-describing objects and tangible data structures

Participants: Nebil Ben Mabrouk, Paul Couderc [contact].

A development in the line of the composite objects (see section 3.3) are self-describing objects. While previous works enabled integrity checking over a set of physical objects, these mechanisms were limited in two aspects: expressiveness and autonomy. More precisely, objects support the detection of special conditions (such as a missing element), but not the characterization of these conditions (such as describing the problem, identifying the missing element). Moreover, this compromises the autonomous feature of coupled objects, which would depend on external systems for analysing these special conditions. Self-describing objects are an attempt to overcome these limitations, and to broaden the application perspectives of autonomous RFID systems.

The principle is to implement distributed data structure over a set of RFID tags, enabling a complex object (made of various parts) or a set of objects belonging to a given logical group to "self-describe" itself and the relation between the various physical elements. Some applications examples includes waste management, assembling and repair assistance, prevention of hazards in situations where various products / materials are combined etc. The key property of self-describing objects is, like for coupled objects, that the vital data are self-hosted by the physical element themselves (typically in RFID chips), not an external infrastructure like most RFID systems. This property provides the same advantages as in coupled objects, namely high scalability, easy deployment (no interoperability dependence/interference), and limited risk for privacy.

However, given the extreme storage limitation of RFID chips, designing such systems is difficult:

- Data structures must be very frugal in terms of space requirements, both for the structure and for the coding.
- Data structures must be robust and able to survive missing or corrupted elements if we want to ensure the self-describing property for a damaged or incorrect object.

In the context of RFID system, the resiliency property of such data structures enables new information architecture and autonomous (offline) operation, which is very important for some RFID applications. We previously applied the self-describing objects approach to the waste management domain [1], which has shown to be a specially challenging situation for RFID. This challenge is found more generally in pervasive computing scenarios involving RFID reading in uncontrolled environments (see section 4.3).

Pervasive support for RFIDs.

We propose to apply our approach to improve the robustness of RFID inventories / batch checking: when many objects are read at once by an RFID reader, miss read are common and raise reliability and operational issues for applications. An innovative solution to this problem is to take advantage of the multiplicity of tags by leveraging them as a distributed memory shared by a logical group. In this way, it is possible to support error detection as well as information recovery. We proposed a flexible protocol to support robust EPC retrieval in adverse reading conditions. The proposed protocol uses erasure correcting techniques to enable error-free recovery of misread EPCs [2]. It is further customizable with respect to the rate of misread tags and application requirements. This work was the object of an Inria patent ⁰. Fine-tuning the protocol parameters is still the object of on going experimentation in the context of the Pervasive_RFID project (see section 7.1.1).

⁰Patent filed in April 2015 - Inria 179

At the software level, RFID inventory reliability issue is usually addressed by anti-collisions mechanisms and redundancy mechanisms. Anti-collisions protocols limit the risk of data corruption when multiples tags have to reply to an inventory request. Redundancy is often implemented in RFID readers by aggregating the results of multiple inventory requests over a time frame, to give the tags multiple opportunities to reply. While useful, these strategies cannot ensure that a given inventory is valid or not (in other words, one or more tags may be missing without being noticed). In situations where we have to read large collection of objects of various types, the performance is difficult to predict but may still be adequate for a given application. For example, some application can tolerate missing some tags, provided that miss read probability could be characterized. In some cases, read reliability could be improved using mechanical approaches, such as introducing movements in objects or antenna to introduce *radio diversity* during read. Finally, distributed data structure can be used over a set of tags to be used to mitigate the impact of misread (by using data redundancy) and to help the reading protocol by integrating hints about the tag set collection being read.

We studied extensively by experimentation the behaviour of existing RFID solutions in the context of uncontrolled environment (meaning, random placement of tags on objects mixing various materials) in order to characterize their real-world performance regarding the parameters of such as tags numbers, density, frequencies, reader antenna design, dynamicity of objects (movements), etc. From these experimentations, we would like to identify the conditions that are favorable to acceptable performance, and the way where there are hopes of improvement with specific design for these difficult environments. These results should also allow improving the performance: high level integrity checks can guide low level operations by determining whether inventories are complete or not. This cross layer strategy enables faster and more efficient inventory protocols.

6.2. Interactions between connected objects in a Smart Building

Participants: Adrien Capaine, Yoann Maurel, Frédéric Weis [contact].

Tacoma group is focussed on the conception and implementation of innovative services for the Smart Home/Building. The range of considered services is broad: from "optimizing the energy consumption" to "helping users to find their way in a building". One of our goals is to build a pervasive platform with constrained performance and cost [7], without disrupting existing spaces. Within this idea, we explored in 2015 the services provided by different modes of interaction in a physical space between neighboring objects, and also between an object and a nearby user.

More precisely, we conducted some experiments with LEDs. Instrumented via a short distance radio interface, a lighting device becomes an unobtrusive connected object that is easy to integrate to a mesh network. A relevant aspect of this platform is the consideration of potential conflicts in data access offered by the connected objects. One of the first scenarios we considered is to operate an LED-based light path to guide the evacuation of a building in the case of a fire alarm. When our objective is to multiply the uses of LED devices ("go beyond lighting", see section 7.1.2), the question is then the priority of access to resources offered by the platform distributed in the environment. Specifically, we addressed the following issues (similar to some of the issues presented in section 3.2):

- How to prioritize the lighting functions (classic) and occasional (but priority) uses of the LED to help in the care of a fire alarm?
- How do you prioritize access to the objects and/or resources that carry these items?

6.3. Context computing for Smart Home

Participants: Yoann Maurel, Frédéric Weis [contact].

To provide services for smart Homes, automation based on pre-set scenarios is ineffective: human behavior is hardly predictable and application should be able to adapt their behavior at runtime depending on the context. We focused on recognizing user's activities to adapt applications behaviours. Our aim is to compute small pieces of context we called *context attributes*. Those context attributes are diverse, for example a presence in a room, the number of people in a room etc.

Building efficient and accurate context information using inexpensive and non-invasive sensors was and is still a great challenge 3.1 . We proved, through the use of dedicated algorithms and a layered architecture that it is achievable when the targeted Home is known - due to the specific and non automated calibration process we used. Among all the available theories, we used the Belief Function Theory (BFT) [8] [9] as it allows to express **uncertainty** and **imprecision**.

Context is computed by a chain a tasks as illustrated in figure 5 :

- The transition between a raw sensor value and a belief function is made through the use of a belief model which maps a sensor value to a belief function. A belief function represents the degree of belief associated to each possible value of the context attribute.
- Then a set of belief functions (corresponding to a set of sensors) can be combined (fused).
- Finally the system can decide what is the "best" value for the context attribute.

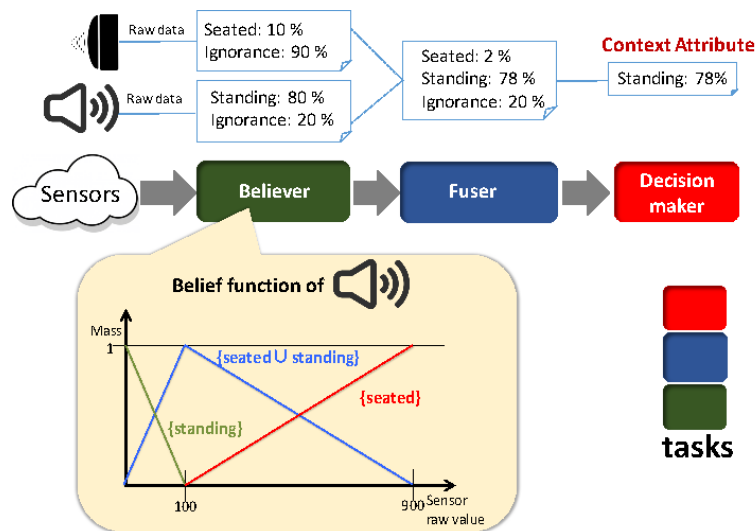


Figure 5. Context attribute computation

Alleviating the complexity of the platform configuration and maintenance is a prerequisite for the adoption of Smart-Home environments by consumers. Currently the BFT theories requires a huge calibration process. We focussed our efforts on the semi-automated building of belief functions, required by the theory, that have to be provided by each sensor.

Automated configuration of sensors.

The belief model is provided to the platform by us and a component is in charge of transforming a sensor value in a *belief function*. The fine tuning of a belief function can be a tedious task. It must be done by a specialist who understands the belief function theory and knows the behavior of the sensors. The model is often built iteratively by experimenting. This may take several hours or days. Moreover, this method is directly connected to the output of each sensor. Biased and noisy measures can cause major modifications on the resulting beliefs.

Ideally, the calibration of the model should be as automatic as possible (few interaction with the user during calibration). The person setting up the sensors should not have to understand the belief function theory. We proposed to generate our belief model from a training set of sensor data. We mainly focused on k-nearest neighbors (KNN) algorithm [6]. We used a training data set to compute the presence belief model. We acquired a set of data with someone present in the experimentation room and a second data set with nobody in the room, which gives us a labelled data set. This principle is illustrated in figure 6 .

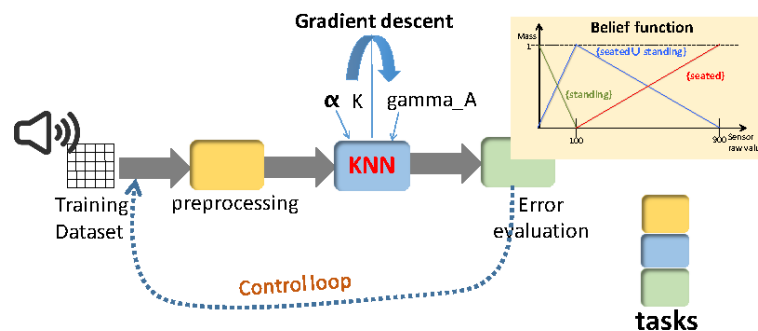


Figure 6. Sensor autocalibration

6.4. Design of a framework for distributed pervasive environments

Participants: Adrien Capaine, Yoann Maurel [contact], Frédéric Weis.

Pervasive environment brings into play complex interactions between a large number of heterogeneous entities: computing units executing third-party applications delivering multiple services to users, with various (sometimes conflicting) requirements, based on the information provided by dynamically (un)available smart object or sensors. The development of pervasive application is consequently hard and must be supported by architectures and frameworks that propose solution to manage the heterogeneity, to organize the interaction of distributed entities, to support the dynamic discovery of the entities, to ensure the privacy of collected data and inferred context, to organize and structure information sharing, and to enforce access control over data and entities.

To alleviate the development of such application (see section 3.4), we worked on a distributed pervasive environment made of several processing nodes (or gateway) managing interacting Smart Spaces (*i.e.* a room, a corridor etc.). A Smart Space contains one or more nodes that coordinate to provide services to users. A node is a low cost computing unit with constrained performances. Each node is responsible for the management of entities and services available in their close proximity: they dynamically discover available devices and source of information, computes contextual information and offer services to nearby users. Nodes are organized hierarchically: in each space one *supervisor node* is responsible for coordination between nodes (*e.g.* managing conflicting requirements and enforcing global policies) and communication with neighboring Smart Spaces. The whole environment (a Smart Building or a Smart Home) is controlled by a master node that distributes policy between Smart Spaces to provide global services (*e.g.* global energy management). Data are stored and processed by nodes as close as possible to the users in order to enforce data privacy (see section 2.1).

Each node supports the execution of several services and application. To help the development of these services, applications are built on top of a framework (**Matriona**) running on each gateways.

Matriona proposes a unified representation of the concepts manipulated by applications in order to hide the heterogeneity of technologies to the application. We rely on the concept of *resource*. A Matriona's resource is quite similar to a REST resource but is not identical: HTTP protocol is only used for remote call; the structure of a resource is constrained; a resource can be dynamically discovered; it can provide notification (PUSH operation); it is uniquely identified in the environment; it is any object used or shared by applications, by bridges, or by the system itself: a device, a room, a platform, a user or a contextual information; it is implemented as a standard object and has a type specified by its interfaces (annotated Java interface). Types describe the data and the operations on resources. *CRUD* (Create, Read, Update, Delete) and *PUSH* operations are used to represent the semantic of operations. Specifying the semantic of operations allows to operate some

generic processing (*e.g.* gathering all the information provided by a resource) on resources without knowing their types.

Matriona enables dynamic discovery between nodes and inter-platform communication between the applications and resources. The platform automatically discovers other platforms using a discovery mechanism. Each platform enables the discovery and the use of its resources to other platforms' applications. Remote resources are using exactly the same API as their internal counterparts: using remote resources is completely transparent to the application. The data is serialized / deserialized automatically by the platform during the call. Calling remote resources induces a performance cost equivalent to the use of a traditional REST resource.

Matriona allows to dynamically add new properties and new behaviors to a resource using decorators. Resources are built using multiple layers in the same way as "Russian dolls". Each layer is responsible for implementing a specific behavior such as retrieving, conversion operation or adding new properties (*e.g.* localization). For instance, the implementation of a thermometer resource will consist of 1) a core layer providing standard information (id, date of creation, the groups to which it belongs); 2) a protocol layer able to communicate with the real devices; 3) a conversion layer (Kelvin to Celsius). The application interacts with the top layer that exposes all or part of information and treatments offered by lower layers. While some layers are static (part of the resource declaration) and cannot be removed, most can be added afterward by the applications. It creates a virtual resource composed of the original resource and the new layers. This virtual resource can be used and discovered as any other resources.

Matriona allows to organize the information. Resource may reference other resources, for example the localization of a "thermometer resource" refers to the resource representing the room in which it is located. The value of the property is the *id* of the referenced resource. This allows applications to easily find resources and their interactions. It is also possible to create composite resources using aggregation mechanisms provided by the framework. The virtual resource can be used directly by applications as any other resources.

Matriona provides a basic language queries. Applications use resources directly or send queries to the platform registry. The query language allows to apply filters and to aggregate data on available resources. A request is represented by a specific URL. For instance, the mean temperature of thermometers in the whole meeting rooms of a building can be obtained using the URL `/**/*.location/meeting_room/temperature!mean`. The query language can be extended by providing new decorators and new filters.

Matriona provides access management: each resources belongs to one or many groups. The groups are defined when the resource is created or during its lifecycle by the owner of the resource. Groups gather applications that share the same permissions on access resources. Groups are managed by a "group owner" that can limit members permissions. Permissions describe the ability of an application to read, write, update, delete, manage or lock a resource. Resource locking avoids conflicting requests to be performed by different applications. Locks are given to applications for a fixed period of time. A resource can always be unlocked by the platform itself or by "critical" applications (*e.g.* emergency fire alarm, see section 6.2).

Matriona allows applications to extend resource properties and to share these meta-information with others. Each application can add new information to a given resource. Tagged information are available only for the application and its group. Meta-information are stored by the platform and associated to the resource until the latter is destroyed. This mechanism allows application to easily share information on the resource they used. For instance, this can be used to retrieve previously used resources or to rate the quality of service provided by a given resource. For the application, meta-information are part of the resource. It is then possible for an application to only use resources that have been approved by other applications of their group. This mechanism can also be used by application to add some task-relevant information (*e.g.* a medical application can tag resources that have been used by a patient).

6.5. Towards Metamorphic Housing: the on-demand room

Participant: Michele Dominici [contact].

This research activity is supported by Fondation Rennes 1 through the chair "Smart Home and Innovation", since January 2014. This activity is centered on the concept of metamorphic housing (see section 4.2). During the first year, we had identified the goals of the research project, also taking into account the trends of future housing industry, provided by the enterprises and public authorities that support the chair. We also had identified a case study, the on-demand room, to be displayed as the main application of the research results in scientific communications and vulgarization. It consists in a space that is physically shared by a small group of apartments, but is assigned for the sole use of one or few particular ones at the time. The room is designed so as to make occupants feel they did not leave their apartment at all. They seamlessly move from their dwelling to the on-demand room and conversely, without noticing the difference, as the room adapts to their preferences. During 2015, second year of the chair, we organized our work following two main axes: (i) solving the research problems, illustrated in the rest of this section; (ii) demonstrating the results using mixed reality as combination of virtual reality and off-the-shelf domotic devices, described in section 5.2.2 .

The research problems underlying the on-demand room are numerous: we illustrated them in the research report "A Case Study of Metamorphic Housing: The on-Demand Room" [3]. We started by addressing the problems associated with the goal of "plugging" the room into different apartments. This requires to dynamically change the rights to control and customize the room's equipment, including lights, appliances, heating, ventilation and air conditioning systems (HVAC), etc. This must be done in a transparent fashion, so that off-the-shelf devices and appliances can be used.

To solve these problems, we started a collaboration with the DIVERSE team ⁰. The goal is to use the Kevoree ⁰ software framework to dynamically reconfigure the networks and domotic system of the room and of the apartments. When the on-demand room is owned by an apartment, their computer networks are interconnected; appliances, sensors and controllers in the room and the apartment can communicate with each others; devices reflect user preferences. Kevoree will enable these reconfigurations by running on key appliances and dynamically adapting and customizing their behavior to the owner of the on-demand room.

As part of the collaboration, some research goals have already been identified. The underlying challenges will be addressed and the results will be integrated in a comprehensive mixed reality demonstrator. This will represent the final iteration of the ongoing demonstration process, illustrated in the platform section (for more details, see 5.2.2).

⁰<http://diverse.irisa.fr/>

⁰<http://kevoree.org/>

TADAAM Team

6. New Results

6.1. TreeMatch Development

This year we have modified the TreeMatch API in order to enable better integration inside application with higher-level abstractions more precise semantic. We also introduced the “over subscribing” features that allow to map more than one process on a given processing unit. We also added new metrics to measure the performance of the proposed placement. We now have three metrics: The sum of the communication cost, the maximum of the communication cost and the hope-byte.

6.2. Affinity Abstraction

This year we worked on the affinity abstraction. Often, the affinity between two processes or threads is measured by the a matrix where a high entry represent a high affinity. Such example of matrices are the number of messages and the number of bytes exchanged between processing units. However, such matrix hide many characteristics of the application such as computation/communication overlap, network contention, etc. First, we have developed a new OpenMPI PML module to gather communication matrix of a running application. Then, we have conducted an extensive study of the minighost application to understand how such communication matrix actually measure the affinity between processes. On this application it appears that the size metric better matches the performances and that the performance of process placement is highly correlated to the proportion of communication in the application.

6.3. Locality for Application Using Locks on Clusters of Multicore Platforms

The aim of this post-doc work is to study the locality for applications based on read-write locks on clusters of multi-core platforms. We focused on the implementation of the video tracking application [25] using the Ordered Read Write Locks (ORWL) [20] model of programming on multi-cores architecture. For several uses, such as, human-computer interaction, security or traffic control, the tracking application aim to locate multiple moving objects over time using a camera. Its processing can be a time consuming process due to the amount of data that is contained in high definition video which leads to decrease the throughput. To overcome this problem it is possible to speed up the processing by exploiting task parallelism of ORWL model. Indeed, the model proposes abstractions of the decomposition in parallel parts (tasks), the synchronization of and the communication between threads. However, we noted some problems which decrease the parallelism scaling thus we introduced different optimizations: stream multiplexing, multiple buffering, etc. We are now working on parallelizing long-running tasks.

6.4. Topology Aware Malleability of MPI programs

Current parallel environments aggregate large numbers of computational resources with a high rate of change in their availability and load conditions. In order to obtain the best performance in this type of infrastructures, parallel applications must be able to adapt to these changing conditions.

In collaboration with Universidade da Coruña, Spain, we have worked on automatically and transparently adapt MPI applications to available resources is proposed. The solution relies on application-level migration approach, incorporating a new scheduling algorithm, based on TreeMatch and Hwloc, to obtain well balanced nodes while preserving performance as much as possible.

The experimental evaluation shows successful and efficient operation, with an overhead of less than 1 second for the proposed scheduling algorithm, and of only a few seconds for the complete reconfiguration, which will be negligible in large applications with a realistic reconfiguration frequency.

6.5. Topology Aware Load Balancing

Charm++ is a message-passing based programming environment that uses an object-oriented approach. However, where MPI considers processes in its model, Charm++ uses finer-grain migratable objects called chares. Brought together with an adaptive runtime system, Charm++ allows to perform dynamic load balancing considering the CPU load of each chare. Our work on data locality and process placement lead us to add the benefits of our TreeMatch algorithm in a load balancing solution. Thus we developed few months ago a topology-aware load-balancer in Charm++ using TreeMatch to reduce the communication costs. During the last months, we significantly improved this load-balancer and its scalability. Particularly, our load balancing algorithm is now hierarchical and distributed. To validate this approach, we have begun to carry out experiments with a cosmological application on the Blue Waters supercomputer. The results will be published soon.

6.6. Topology Aware Resource Management

SLURM [24] is a Resource and Job Management System, a middleware in charge of delivering computing power to applications in HPC systems. Our goal is to take in account in SLURM placement process hardware topology as well as application communication pattern. We have a new selection option for the `cons_res` plugin in SLURM. In this case the usually BestFit algorithm used to choose nodes is replaced by TreeMatch to find the best placement among the free nodes list in light of a given application communication matrix.

We updated this plugin based on SLURM 2.6.5 for last version SLURM 15.08. To decrease the overhead due to our algorithm we also implemented an alternative to use a subtree of the global topology. We ran experiments to compare these different solutions using our plugin with or without subtree and the current algorithm topology-aware in SLURM.

6.7. Topology Aware Performance Monitoring

While system's scale is growing exponentially, memory hierarchy is getting larger, at various levels. Hence optimizing applications to reach an optimal usage of a machine may involve a large spectrum of performance metrics interacting at different level of the system's hierarchy. Memory bound applications showing irregular patterns lead to locality issues. Addressing those issues and getting a good schedule on complex systems is a NP hard problem and can therefore only be solved with heuristics. Although powerful algorithms using the most intuitive heuristics such as communications path reduction and/or cache contention reduction may show good results on some cases, there are still room for improvements in this direction so much the configuration of applications, systems, software stack vary and impact the execution time.

In order to step in this direction we developed a highly extendible tool to gather asynchronously performance data from different sources. This information is then aggregated into different topology objects (cache, node, processing unit, ...) in order to give a synthetic and topology aware information to drive optimization.

In brief the tool works this way: The user provide a description file with arithmetic expression of performance counters(defined into performance data plugins), and topology objects where to map the expression. A pair (expression,object) defines a monitor which will sample performance data and stored them into an history. Then others monitors can be defined as a combination of the previous. For instance we can attach a process and record on each core its L3 cache miss counter, and then add each of those monitor into an upper monitor located on the L3 cache. Several aggregation functions are already available but we aim to provide several statistical function to extend the possibility of data interpretation. Such functions allow to aggregate results in a meaningful way. Then we add a locality insight using `lstopo` tool from `hwloc` to draw the results on a topology. This has been published in [12]

6.8. Memory Hierarchy Aware Roofline Model

The increasing complexity of computer architectures, makes challenging to fully exploit computer systems' capabilities. The cost of tuning applications on such machines can raise quickly. Therefore, linking the information about a machine performance bounds and applications performance results respectively to those bounds can help finding the bottlenecks and motivate code optimization.

In 2009 the Roofline model [23] throws those bases by plotting on a 2 dimensional diagram, application performance (GFlop/s) and arithmetic intensity (Flop/Byte) with respect to the main memory bandwidth (GByte/s) and peak floating point performance (GFlop/s). In 2014 the model extended by Alexandar Illic, take into account the data movement inside the cache hierarchy to provide a finer analyse by showing application's performance results with respect to the different cache bandwidths.

With the cooperation of the Cache Aware Roofline Model authors, we have worked on extending this model to the whole memory hierarchy at NUMA scale in order to drive optimisations on next generation processors embedding different memory technologies and different memory configurations like Intel's KNL does.

While we are designing a tool based on hwloc and micro-kernels to empirically extract and validate machines bottlenecks, we also want to show with real NUMA applications that the model may be extended to such hierarchy levels, still providing insightful representation.

6.9. Topology Management and Standardization

We continued to work on the diffusion of our software and ideas in existing programming interfaces and standards tailored for HPC and parallel computing. In particular, we did integrate our TreeMatch algorithm in the Open MPI implementation of the Message Passing Interface, so as to provide enhanced Virtual Topology routines in MPI allowing the user to effectively create parallel applications taking into account both their behaviour and the characteristics of the underlying hardware. Our code is available in the master repository and should be available in an Open MPI distribution at some point in the next year (2016).

We also drafted and submitted a proposal to modify the MPI interface so that information regarding the underlying physical topology could be made available at the MPI application level. We plan to push our ideas during the next year so that our proposal can eventually make its way into the MPI standard.

6.10. Modeling Next-Generation Memory Architectures

We initiated a research topic on modeling next generation memory architectures that will mix different kinds of memories. Indeed the arrival of high-bandwidth and non-volatile memories cause computing cores to have different local memory banks with different characteristics.

The hwloc software 5.1 is being extended in collaboration with CPU vendors such as Intel and AMD to better represent these new memory technologies. We are working with Bull in the context of Nicolas Denoyelle's PhD on developing abstractions for deciding where to allocate the application buffers.

6.11. Modeling Affinity of Multithreaded Applications

With the increasing complexity and scale of multi-core processors, optimizing thread placement becomes more and more challenging. Our goal is to better understand which characteristics of a multi-threaded application can have an impact on a placement decision for a given architecture. To this end, we analyze the performance of a set of applications under different placement strategies and we try to relate the obtained results to characteristics of the applications such as the data footprint of each thread, the amount of data shared between threads, or the reuse distance.

To collect information about the characteristics of multi-threaded applications, we developed a set of tools based on the PIN dynamic binary instrumentation tools. PIN allows us to get information about all instructions executed and memory location accessed by each thread of an application during its execution, and this without modifying the source code of the application.

We used our PIN-based tools to study a representative set of applications taken from two well-known benchmark suites, namely the Mantevo benchmark suites (HPC applications based on OpenMP) and the Parsec benchmark suites (general-purpose applications based on pthreads). Analyzing the results of all our tests is an ongoing work.

6.12. Thread placement and threads policy on a multicore machine with NUMA effects.

Threads placement on multicore machine with NUMA effects is inevitable to have better performances. Threads must bind on cores to avoid thread migration and to have better cache locality. MPI non-blocking collectives can generate progress threads to complete communications. These additional threads can disturb computational threads. That is why we have implemented several thread placement algorithms into the MPC framework [22]. These algorithms allow to dedicate resources only for progress threads. Thus computational threads are not disturb. We test them with our own benchmarks which test all the MPI non-blocking collectives to compare the performances with different thread placement. We observe an improvement when resources are dedicated to progress threads and take NUMA effects into account.

We want to include a mechanism into MPC to specify thread kinds (MPI, OpenMP,...). These mechanism will allow the MPC scheduler to take threads specificity into account to improve the scheduling policy. Our goal is to increase runtime performances considering each type specific needs. We have begun to implement this mechanism.

Several MPC framework bugs have been corrected, thus we contribute to its stability.

6.13. Multithreaded Communications

To program clusters of multicores, hybrid models mixing MPI+threads, and in particular MPI+OpenMP are gaining popularity. This imposes new requirements on communication libraries, such as the need for MPI_THREAD_MULTIPLE level of multi-threading support. Moreover, the high number of cores brings new opportunities to parallelize communication libraries, so as to have proper background progression of communication and communication/computation overlap.

We have proposed PIOMan [11], a generic framework to be used by MPI implementations, that brings seamless asynchronous progression of communication by opportunistically using available cores. It uses system threads and thus is composable with any runtime system used for multithreading. Through various benchmarks, we demonstrated that our pioman-based MPI implementation exhibits very good properties regarding overlap, progression, and multithreading, and outperforms state-of-art MPI implementations.

6.14. RDMA-based Communications

High-performance network hardware is nowadays dominated by RDMA-oriented technologies. The software stack is moving too towards Remote Memory Access. However, most communication libraries stil use send/receive paradigm as a common denominator. We have proposed to study a software stack for networks the is based on remote memory access from the hardware up to the enduser API, where RDMA is first class citizen and not a compatibility layer. It is expected to obtain better performance, better scalability with regard to number of communication flows or threads, and better asynchronous progression, while optimization strategies on the packet flows such as aggregation as proposed in NewMadeleine are still possible. Work has begun as a Masters thesis [18] and continues as Romain Prou Ph.D. thesis.

6.15. Network Modeling

Netloc is a tool for hwloc [1] to find the topology of a supercomputer. For that, it discovers all the networks by exploring them by using tools specifying to the network type. The exploration gives all the machines and all the switches, with all the links between them. We improved netloc by adding the visualization of the topologies discovered. The visualization is dynamic and the user can interact with it, to get some information about the machines, the switches or the link such as the physical address, the hostname or the speed of the link. In order to be able to do optimizations that can be helping process placement, we started to class the different topologies. For now, we only handle Clos networks [21] and we are able to transform them into fat trees. The categorization in classes permits to have a clean graph and then interact with graph partitioners.

To have a complete tool, we need to handle all major classes of topologies such as meshes, torus or hypercubes. When the graph partitioning will be integrated with tools such as SCOTCH, we will be able to find a good mapping for the processes of a job. It could also help the resource scheduling to optimize the resource sharing between jobs. The visualization can be improved by showing the architecture information retrieved by `hwloc` for each machine. We can complete the visualization by giving more information especially when the original graph was transformed to simplify it, as we did to Clos networks to obtain fat trees.

6.16. Scalable mapping onto (disconnected) parts of regular target architectures

Since its inception, SCOTCH allows one to map graphs onto so-called “algorithmically-defined” target architectures. They are regular architectures such as hypercube, multi-dimensional grids and tori, butterfly networks, etc., whose characteristics are defined by subroutines which are part of the SCOTCH library. However, on today’s large-scale computer systems, software jobs do not usually run on all of the machine, but on a set of nodes assigned by the batch scheduler. Consequently, one should be able to map a process graph onto (possibly disconnected) parts of an algorithmically-defined target architecture, which was not an available feature. Only “decomposition-defined” architectures (another way to represent target architectures in SCOTCH) supported this feature, but are not scalable above a few hundred processing elements.

In order to allow SCOTCH to provide mappings onto parts of an algorithmically-defined target architecture, a new meta target architecture, called “sub”, has been created. The sub architecture allows one to restrict a regular algorithmically-defined target architecture to a subset of its vertices. Instead of using a top-down approach to build a description of the target architecture, through a recursive bipartitioning algorithm, our new algorithm uses a bottom-up approach, based on recursive matching and coarsening of neighboring vertices, much like for graph coarsening. The clustering tree is pruned of branches that lead to parts of the machine that are not allowed mapping targets. Distance between subdomains is computed using the distance function of the underlying algorithmically-defined target architecture. Preliminary results have been presented at a SIAM CS&E conference workshop [14], and a beta-version of the upcoming release 6.0.5 of SCOTCH has been shipped to early testers at Lawrence Livermore National Laboratory.

6.17. Multi-Level Parallelism in a CFD code

Code_Saturne [19] is an industrial and open source Computational Fluid Dynamics software. Developed at EDF R&D, it solves the Navier-Stokes equations for 2D, 2D-axisymmetric and 3D flows, steady or unsteady, laminar or turbulent, incompressible or weakly dilatable, isothermal or not, with scalars transport if required.

Our goal is to evaluate different ways of improving and preparing this application for the future HPC architectures. We strengthened our application knowledge by using various instrumentation tools and provided a small topology instrumentation library. As instrumentation of a full code can be a tedious thing, we provided a mini application on which to perform our future experiments. We have run experiments to determine the potential gain of topology awareness on our code by using the graph mapping solutions of PT-SCOTCH. We have also run experiments on ghost cells numbering to see the impact of their locations on cache misses.

URBANET Team

7. New Results

7.1. Characterizing and deploying urban networks

Participants: Ahmed Boubrima, Angelo Furno, Diala Naboulsi, Patrice Raveneau, Walid Bechkit, Marco Fiore, Hervé Rivano, Razvan Stanica.

7.1.1. Collection and Analysis of Mobile Phone Data

Cellular communications are undergoing significant evolutions in order to accommodate the load generated by increasingly pervasive smart mobile devices. At the same time, recent generations of mobile phones, embedding a wide variety of sensors, have fostered the development of open sensing applications, while cellular operators are looking for new services they can provide using the data collected on their side, in the access or the core network.

The analysis of operator-side data is a recently emerged research field, and, apart a few outliers, relevant works cover the period from 2005 to date, with a sensible densification over the last three years. In [9], we provided a thorough review of the multidisciplinary activities that rely on mobile traffic datasets, identifying major categories and sub-categories in the literature, so as to outline a hierarchical classification of research lines and proposing a complete introductory guide to the research based on mobile traffic analysis. The usage of these datasets in the design of new networking solutions, in order to achieve the so-called *cognitive networking* paradigm, is discussed in detail in the PhD thesis of Diala Naboulsi [2], where the examples of green networking and virtualized radio access networks are given.

When constructing a social network from interactions among people (e.g., phone calls, encounters), a crucial task is to define the threshold that separates social from random (or casual) relationships. The ability to accurately identify social relationships becomes essential to applications that rely on a precise description of human routines, such as recommendation systems, forwarding strategies and opportunistic dissemination protocols. We thus proposed a strategy to analyze users' interactions in dynamic networks where entities act according to their interests and activity dynamics [10]. Our strategy allows classifying users interactions, separating random ties from social ones, and unveils significant differences among the dynamics of users' wireless interactions in the datasets.

Furthermore, mobile traffic data has been recently used to characterize the urban environment in terms of urban fabric profiles. While showing promising results, the existing urban fabric detection solutions are built without a clear understanding of the detection process chain. In [16], we distinguished and analyzed the different steps common to all urban profiling techniques. By evaluating the impact of each step of the process, we were able to propose a new solution that outperforms the state of the art techniques. Our approach uses the weekly periodicity of human activities, as well as a median-based filtering technique, resulting in a better clustering in terms of both coverage and entropy, as shown by results obtained on two large scale mobile traffic datasets covering the urban areas of Milan and Turin, in Italy. The solution proposed in this work was selected among the 10 finalists of the Telecom Italia Big Data challenge.

A second source of mobile data is the smartphone itself. In the context of the PrivaMov project, funded by the Labex IMU, we developed and deployed a data collection platform on more than 100 Android devices. A first step in the study of this enormous dataset (more than 50 Gb have been collected to date) was presented in [21], with a focus on the extraction of user mobility information and Wi-Fi mapping. This led us to the study of Wi-Fi tracking, a method relying on signals emitted by portable devices to track individuals for commercial, security or surveillance purposes. Wi-Fi tracking has the potential to passively track a large fraction of the population and is therefore an ideal population surveillance technology and a serious privacy threat. In [19], we argue that Wi-Fi routers make an ideal building block to create a large scale Wi-Fi tracking system, showing how they can be easily turned into Wi-Fi tracking devices through software modification. We

provided a first evaluation of the tracking capabilities of an hypothetical Wi-Fi tracking system through a set of simulations based on real-world datasets. Results showed that the spatial distribution of Wi-Fi routers is such that compromising even a small fraction of Wi-Fi routers is sufficient to track people for a large fraction of the time.

Preservation of user privacy is therefore paramount in the publication of datasets that contain fine-grained information about individuals. The problem is especially critical in the case of mobile traffic datasets collected by cellular operators, as discussed above, as they feature high subscriber trajectory uniqueness and they are resistant to anonymization through spatiotemporal generalization. In [17], we first unveiled the reasons behind such undesirable features of mobile traffic datasets, by leveraging an original measure of the anonymizability of users' mobile fingerprints. Building on such findings, we proposed GLOVE, an algorithm that grants k -anonymity of trajectories through specialized generalization. We evaluated our methodology on two nationwide mobile traffic datasets, and show that it achieves k -anonymity while preserving a substantial level of accuracy in the data.

7.1.2. Deployment of Wireless Sensor Networks for Pollution Monitoring

Recently, air pollution monitoring emerged as one of the main services of smart cities because of the increasing industrialization and the massive urbanization. Wireless Sensor Networks are a suitable technology for this purpose, thanks to their substantial benefits including low cost and autonomy. Minimizing the deployment cost is one of the major challenges in the design of such networks, therefore sensors positions have to be carefully determined. In [13], we proposed two integer linear programming formulations based on real pollutants dispersion modeling to deal with the minimum cost sensor network deployment for air pollution monitoring. We illustrated the concept by applying our models on real world data, namely the Nottingham City street lights. We compared the two models in terms of execution time and showed that the second flow-based formulation is much better. We finally conducted extensive simulations to study the impact of some parameters and derive some guidelines for efficient urban sensor deployment for air pollution monitoring.

7.2. Technology specific solutions

Participants: Jin Cui, Walid Bechkit, Khaled Boussetta, Hervé Rivano, Fabrice Valois.

7.2.1. Temperature-Aware Algorithms for Wireless Sensor Networks

Temperature variations have a significant effect on low power wireless sensor networks as wireless communication links drastically deteriorate when temperature increases. A reliable deployment should take temperature into account to avoid network connectivity problems resulting from poor wireless links when temperature increases. A good deployment needs also to adapt its operation and save resources when temperature decreases and wireless links improve. Taking into account the probabilistic nature of the wireless communication channel, in [12] we investigated the effect of temperature on percolation-based connectivity in large scale wireless sensor networks and showed that more energy can be saved by allowing some nodes to go to deep sleep mode when temperature decreases and links improve. Based on this result, we proposed a simple, yet efficient, Temperature-Aware MAC plugin (TA-MAC), which can be potentially used with any MAC protocol, enabling it to dynamically adapt the network effective density in order to allow further energy savings, while maintaining network connectivity. We carried out simulations and demonstrated that state of the art protocols augmented with the TA-MAC plugin allow a significant energy efficiency improvement.

Going one step further, we developed a mathematical model that provides the most energy efficient deployment in function of temperature without compromising the correct operation of the network by preserving both connectivity and coverage [3]. We used our model to design three temperature-aware algorithms that seek to save energy (i) by putting some nodes in hibernate mode as in the SO (Stop-Operate) algorithm in TA-MAC, or (ii) by using transmission power control as in PC (Power-Control), or (iii) by doing both techniques as in SOPC (Stop-Operate Power-Control). All proposed algorithms are fully distributed and solely rely on temperature readings without any information exchange between neighbors, which makes them low overhead and robust. Our results identified the optimal operation of each algorithm and showed that a significant amount of energy can be saved by taking temperature into account.

7.2.2. Resilience in Wireless Sensor Networks

The concept of resilience for routing protocols in wireless sensor networks has been proposed and developed in the team in the last few years. In our previous works, a general overview of the resilience, including definition, metric and resilient techniques based on random behavior and data replication have been proposed. Following these previous methods, in [6] we proposed a new resilient solution based on network coding techniques, to improve resilience in wireless sensor networks for smart metering applications. More precisely, using our resilience metric based on a performance surface, we compared several variants of a well-known gradient based routing protocol with the previous methods (random routing and packet replications) and the new proposed methods (two network coding techniques). The proposed methods outperformed the previous methods in terms of data delivery success even in the presence of high attack intensity.

We also continued to study the resilience of routing protocols against malicious insiders willing to disrupt network communications. Previously, the simulation results showed that introducing randomness in routing protocols increases uncertainty for an adversary, making the protocols unpredictable. When combined with data replication, it permits route diversification between a source and a destination, thus enhancing the resilience. In [15], we proposed a theoretical framework to quantify analytically the performance of random protocols against attacks based on biased random walks on a torus lattice. The objective is to evaluate analytically the influence of bias and data replication introduced to random walks. The bias allows to decrease the route length by directing random walks toward the destination, thus reducing the probability of a data packet to meet a malicious insider along the route; however, it decreases also the degree of randomness (entropy). When random protocols are combined with data replication, the reliability is improved thanks to route diversity despite an additional overhead in terms of energy consumption.

7.2.3. Data aggregation in Wireless Sensor Networks

Aggregation functions are intended to save energy and capacity in Wireless Sensor Networks, by avoiding unnecessary transmissions. Aggregation functions take benefit from spatial and/or temporal correlations to forecast or to compress the real data which are collected. Although several works have focused on data aggregation in Wireless Sensor Networks, there is a lack of a formal unified framework that can compare several aggregation functions suitable for a given network topology, a given application and a target accuracy. In [14], we address this question by proposing a Markov Decision Process that can help to evaluate the performances of aggregation functions. The performances are expressed using two new proposed metrics, which can assess the energy and capacity savings of aggregation functions. As illustrative examples, we use our Markov Decision Process to evaluate and analyze the performances of basic aggregation functions (e.g. average) and more complex ones (time series, polynomial functions).

7.2.4. Data Gathering in Mesh Networks

In the gathering problem in mesh networks, a particular node in a graph, the base station, aims at receiving messages from some nodes in the graph. At each step, a node can send one message to one of its neighbors (such an action is called a call). However, a node cannot send and receive a message during the same step. Moreover, the communication is subject to interference constraints, more precisely, two calls interfere in a step, if one sender is at distance at most d_I from the other receiver. Given a graph with a base station and a set of nodes having some messages, the goal of the gathering problem is to compute a schedule of calls for the base station to receive all messages as fast as possible, i.e., minimizing the number of steps (called makespan). The gathering problem is equivalent to the personalized broadcasting problem where the base station has to send messages to some nodes in the graph, with same transmission constraints.

In [5], we focused on the gathering and personalized broadcasting problem in grids. Moreover, we considered the non-buffering model: when a node receives a message at some step, it must transmit it during the next step. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we presented linear (in the number of messages) algorithms that compute schedules for gathering with $d_I \in \{0, 1, 2\}$. In particular, we presented an algorithm that achieves the optimal makespan up to an additive constant 2 when $d_I = 0$. If no messages are “close” to the axes (the base station being the origin),

our algorithms achieve the optimal makespan up to an additive constant 1 when $d_I = 0$, 4 when $d_I = 2$, and 3 when both $d_I = 1$ and the base station is in a corner.

7.3. Capillary Network Solutions

Participants: Patrice Raveneau, Trista Lin, Marco Fiore, Hervé Rivano, Razvan Stanica.

7.3.1. Connected Vehicles

Managing user mobility is historically one of the most critical issues in cellular radio access networks (RANs). That task will become an even greater challenge due to cellular users on-board vehicles and networked cars that autonomously access Internet-based services, whose number is expected to grow dramatically in the next few years. There is thus a need to characterize RAN access from/by vehicles in a similar way to what has been done for traditional pedestrian access. In [11], we proposed a first study of the macroscopic and microscopic features of pervasive vehicular access in a case-study large-scale urban environment, in presence of realistic datasets of the road traffic and RAN deployment. We found that pervasive vehicular access is characterized by unique temporal and spatial variability in the urban region, such that it may require a dedicated RAN capacity planning: the presence of stable vehicular access load patterns and mobility flows can help to that end. Also, we identified the theoretical distributions that best fit key metrics for RAN planning, i.e., the vehicular users' inter-arrival and residence times at cells, and discuss how their parameters vary over time and space.

Smart parking, allowing drivers to access parking information through their smart-phone, is another important service for vehicular users, which can be provided not only through cellular networks, but also by using metropolitan wireless networks, whose deployment strategy needs to be guided by efficiency and functionality. In [8], we introduced and studied a deployment strategy for wireless on-street parking sensor networks. We defined a multiple-objective problem in our analysis, and solved it with two real-world street parking maps. We presented the results on the tradeoff among minimum energy consumption, sensing information delay and the amount of deployed mesh routers and Internet gateways, i.e., the cost of city infrastructure. We also analyzed these tradeoffs to see how different urban layouts affect the optimal solutions. The overall smart parking architecture and services made the object of the PhD thesis of Trista Lin [1], where the analysis of the entire system can be found, including results on the wireless sensor networks used to collect data from parking places and the Publish-Subscribe service used to disseminate this information to users.

7.3.2. Offloading Cellular Networks

Offloading is a promising technique for alleviating the ever-growing traffic load from infrastructure-based networks such as the Internet. Offloading consists in using alternative methods of transmission as a cost-effective solution for network operators to extend their transport capacity. Wi-Fi offloading is one of the most effective approaches to relieve the cellular radio access from part of the burgeoning mobile demand. To date, Wi-Fi offloading has been mainly leveraged in limited contexts, such as home, office or campus environments. In [18], we investigated the scaling properties of Wi-Fi offloading, by studying how it would perform on a much larger scope than those considered today. To that end, we considered a real-world citywide scenario, built on data about actual infrastructure deployments and mobile traffic demand, and observed which amount of traffic could be accommodated by the existing pervasive Wi-Fi access infrastructure, were it opened to mobile users. We found that more than 80% of the mobile traffic demand in a large urban area may be easily served by Wi-Fi access points, under a wide range of system settings.

A new offloading technique was introduced in [20] and further detailed in [4], where we advocate the use of conventional vehicles equipped with storage devices as data carriers whilst being driven for daily routine journeys. The road network can be turned into a large-capacity transmission system to offload bulk transfers of delay-tolerant data from the Internet. The challenges we addressed include how to assign data to flows of vehicles and while coping with the complexity of the road network. We proposed an embedding algorithm that computes an offloading overlay where each logical link spans over multiple stretches of road from the underlying road infrastructure. We then formulated the data transfer assignment problem as a novel linear programming model we solve to determine the optimal logical paths matching the performance requirements

of a data transfer. We evaluated our road traffic allocation scheme using actual road traffic counts in France. The numerical results show that 20% of vehicles in circulation in France equipped with only one Terabyte of storage can offload Petabyte transfers in a week.

WHISPER Project-Team

7. New Results

7.1. Software engineering for infrastructure software

Tracking code fragments of interest is important in monitoring a software project over multiple versions. Various approaches, including our previous work on Herodotos, exploit the notion of Longest Common Subsequence, as computed by readily available tools such as GNU Diff, to map corresponding code fragments. Nevertheless, the efficient code differencing algorithms are typically line-based or word-based, and thus do not report changes at the level of language constructs. Furthermore, they identify only additions and removals, but not the moving of a block of code from one part of a file to another. Code fragments of interest that fall within the added and removed regions of code have to be manually correlated across versions, which is tedious and error-prone. When studying a very large code base over a long time, the number of manual correlations can become an obstacle to the success of a study. In a paper published at the IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER) [14], we investigate the effect of replacing the current line-based algorithm used by Herodotos by tree-matching, as provided by the algorithm of the differencing tool GumTree. In contrast to the line-based approach, the tree-based approach does not generate any manual correlations, but it incurs a high execution time. To address the problem, we propose a hybrid strategy that gives the best of both approaches.

Understanding the severity of reported bugs is important in both research and practice. In particular, a number of recently proposed mining-based software engineering techniques predict bug severity, bug report quality, and bug-fix time, according to this information. Many bug tracking systems provide a field "severity" offering options such as "severe", "normal", and "minor", with "normal" as the default. However, there is a widespread perception that for many bug reports the label "normal" may not reflect the actual severity, because reporters may overlook setting the severity or may not feel confident enough to do so. In many cases, researchers ignore "normal" bug reports, and thus overlook a large percentage of the reports provided. On the other hand, treating them all together risks mixing reports that have very diverse properties. In a study published at the Working Conference on Mining Software Repositories (MSR) 2015 [16], we investigate the extent to which "normal" bug reports actually have the "normal" severity. We find that many "normal" bug reports in practice are not normal. Furthermore, this misclassification can have a significant impact on the accuracy of mining-based tools and studies that rely on bug report severity information.

Software is continually evolving, to fix bugs and add new features. Industry users, however, often value stability, and thus may not be able to update their code base to the latest versions. This raises the need to selectively backport new features to older software versions. Traditionally, backporting has been done by cluttering the backported code with preprocessor directives, to replace behaviors that are unsupported in an earlier version by appropriate workarounds. This approach however involves writing a lot of error-prone backporting code, and results in implementations that are hard to read and maintain. In a paper published at the 2015 European Dependable Computing Conference (EDCC) [15], we consider this issue in the context of the Linux kernel, for which older versions are in wide use. We present a new backporting strategy that relies on the use of a backporting compatibility library and on code that is automatically generated using the program transformation tool Coccinelle. This approach reduces the amount of code that must be manually written, and thus can help the Linux kernel backporting effort scale while maintaining the dependability of the backporting process.

Logging is a common and important programming practice, but choosing how to log is challenging, especially in a large, evolving software code base that provides many logging alternatives. Insufficient logging may complicate debugging, while logging incorrectly may result in excessive performance overhead and an overload of trivial logs. The Linux kernel has over 13 million lines of code, over 1100 different logging functions, and the strategies for when and how to log have evolved over time. To help developers log correctly

we propose, in a paper published at BEneVol 2015 [18], a framework that will learn existing logging practices from the software development history, and that will be capable of identifying new logging strategies, even when the new strategies just start to be adopted.

7.2. Java runtime support

Java class loaders are commonly used in application servers to load, unload and update a set of classes as a unit. However, unloading or updating a class loader can introduce stale references to the objects of the outdated class loader. A stale reference leads to a memory leak and, for an update, to an inconsistency between the outdated classes and their replacements. To detect and eliminate stale references, in a paper published at DSN 2015 [12], we propose Incinerator, a Java virtual machine extension that introduces the notion of an outdated class loader. Incinerator detects stale references and sets them to null during a garbage collection cycle. We evaluate Incinerator in the context of the OSGi framework and show that Incinerator correctly detects and eliminates stale references, including a bug in Knopflerfish. We also evaluate the performance of Incinerator with the DaCapo benchmark on VMKit and show that Incinerator has an overhead of at most 3.3%.

7.3. Parallel and Distributed Computing

The scalability of multithreaded applications on current multicore systems is hampered by the performance of lock algorithms, due to the costs of access contention and cache misses. In an article published in ACM Transactions on Computer Systems (TOCS), we present a new locking technique, Remote Core Locking (RCL) [10], that aims to accelerate the execution of critical sections in legacy applications on multicore architectures. The idea of RCL is to replace lock acquisitions by optimized remote procedure calls to a dedicated server hardware thread. RCL limits the performance collapse observed with other lock algorithms when many threads try to acquire a lock concurrently and removes the need to transfer lock-protected shared data to the hardware thread acquiring the lock because such data can typically remain in the server's cache. Other contributions presented in this article include a profiler that identifies the locks that are the bottlenecks in multithreaded applications and that can thus benefit from RCL, and a reengineering tool that transforms POSIX lock acquisitions into RCL locks. Eighteen applications were used to evaluate RCL: the nine applications of the SPLASH-2 benchmark suite, the seven applications of the Phoenix 2 benchmark suite, Memcached, and Berkeley DB with a TPC-C client. Eight of these applications are unable to scale because of locks and benefit from RCL on an x86 machine with four AMD Opteron processors and 48 hardware threads. By using RCL instead of Linux POSIX locks, performance is improved by up to 2.5 times on Memcached, and up to 11.6 times on Berkeley DB with the TPC-C client. On a SPARC machine with two Sun Ultrasparc T2+ processors and 128 hardware threads, three applications benefit from RCL. In particular, performance is improved by up to 1.3 times with respect to Solaris POSIX locks on Memcached, and up to 7.9 times on Berkeley DB with the TPC-C client.

Software Transactional Memory (STM) is an optimistic concurrency control mechanism that simplifies parallel programming. Still, there has been little interest in its applicability for reactive applications in which there is a required response time for certain operations. In an article published in ACM Transactions on Parallel Computing (TOPC) [11], we propose supporting such applications by allowing programmers to associate time with atomic blocks in the forms of deadlines and QoS requirements. Based on statistics of past executions, we adjust the execution mode of transactions by decreasing the level of optimism as the deadline approaches. In the presence of concurrent deadlines, we propose different conflict resolution policies. Execution mode switching mechanisms allow meeting multiple deadlines in a consistent manner, with potential QoS degradations being split fairly among several threads as contention increases, and avoiding starvation. Our implementation consists of extensions to a STM runtime that allow gathering statistics and switching execution modes. We also propose novel contention managers adapted to transactional workloads subject to deadlines. The experimental evaluation shows that our approaches significantly improve the likelihood of a transaction meeting its deadline and QoS requirement, even in cases where progress is hampered by conflicts and other concurrent transactions with deadlines.

A challenge in designing a peer-to-peer (P2P) system is to ensure that the system is able to tolerate selfish nodes that strategically deviate from their specification whenever doing so is convenient. In a paper published at SRDS 2015 [13], we propose *RACOON*, a framework for the design of P2P systems that are resilient to selfish behaviours. While most existing solutions target specific systems or types of selfishness, *RACOON* proposes a generic and semi-automatic approach that achieves robust and reusable results. Also, *RACOON* supports the system designer in the performance-oriented tuning of the system, by proposing a novel approach that combines Game Theory and simulations. We illustrate the benefits of using *RACOON* by designing two P2P systems: a live streaming and an anonymous communication system. In simulations and a real deployment of the two applications on a testbed comprising 100 nodes, the systems designed using *RACOON* achieve both resilience to selfish nodes and high performance.

7.4. From Sets to Bits in Coq

Sets form the building block of mathematics, while finite sets are a fundamental data structure of computer science. In the world of mathematics, finite sets enjoy appealing mathematical properties, such as a proof-irrelevant equality and extensionality of functions. Computer scientists, on the other hand, have devised efficient algorithms for set operations based on the representation of finite sets as bit vectors and on bit twiddling, exploiting the hardware's ability to efficiently process machine words.

With interactive theorem provers, sets are reinstated as mathematical objects. While there are several finite set libraries in COQ, these implementations are far removed from those used in efficient code. Recent work on modeling low-level architectures, such as x86 [41] processors, however, have brought the world of bit twiddling within reach of our proof assistants. We are now able to specify and reason about low-level programs.

In this work, we have implemented bitsets and their associated operations in the Coq proof assistant, thus allowing us to transparently navigate between the concrete world of bit vectors and the abstract world of finite sets. This work grew from a puzzled look at the first page of Warren's *Hacker's Delight* [77], where lies the cryptic formula $x \& (x - 1)$ to turn off the rightmost bit in a word. How do we translate the English specification given in the book into a formal definition? How do we prove that this formula meets its specification? Could COQ generate efficient and trustworthy code from it? And how efficiently could we simulate it within COQ itself?

In our work, we have established a bijection between bitsets and sets over finite types. Following a refinement approach, we have shown that a significant part of `SSREFLECTfinset` library can be refined to operations manipulating bitsets. We have also developed a trustworthy extraction of bitsets down to OCaml's machine integers. While we were bound to axiomatize machine integers, we adopted a methodology based on exhaustive testing to gain greater confidence in our model. Finally, we have demonstrated the usefulness of our library through two applications, a certified implementation of Bloom filters and a verified implementation of the n -queens algorithm.

ALICE Project-Team

7. New Results

7.1. Dihedral Angle-Based Maps of Tetrahedral Meshes

Participants: Nicolas Ray, Bruno Lévy.

This work is a collaboration with Gilles-Philippe Paillé (visiting), Pierre Poulin (U. de Montréal) and Alla Sheffer (UBC).

Given a 2D triangulation, it is well known that it is reasonably easy to reconstruct the shape of all the triangles from the sole data of the angles at the triangle corners, provided that they satisfy some constraints. In this project, we studied how this idea can be generalized in the volumetric setting. In other words, we proposed a geometric representation of a tetrahedral mesh that is solely based on dihedral angles, and what are the constraints that these dihedral angles need to satisfy to make that possible. We first show that the shape of a tetrahedral mesh is completely defined by its dihedral angles. This proof leads to a set of angular constraints that must be satisfied for an immersion to exist in \mathbb{R}^3 . This formulation lets us easily specify conditions to avoid inverted tetrahedra and multiply-covered vertices, thus leading to locally injective maps. We then present a constrained optimization method that modifies input angles when they do not satisfy constraints. Additionally, we develop a fast spectral reconstruction method to robustly recover positions from dihedral angles. We demonstrate the applicability of our representation with examples of volume parameterization, shape interpolation, mesh optimization, connectivity shapes, and mesh compression. This work has been published in Transactions on Graphics [17].

7.2. Hexahedral-dominant Remeshing

Participants: Dmitry Sokolov, Nicolas Ray, Bruno Lévy, Maxence Reberol.

Representing the geometry of complex objects in a computer is usually achieved by a mesh: the object is decomposed in cells that have a simple geometry. Each cell is defined by a set of facets. The simplest choice is to use meshes with tetrahedral cells that are relatively easy to produce and to work with. However, some applications involving numerical simulations better work with hexahedral cells. Such hexahedral meshes are very difficult to produce, even when it is completely done by a designer.

Our objective is to relax the intrinsic difficulties of full hexahedral remeshing by allowing the process to generate a few tetrahedra in the hexahedral mesh (hexahedral-dominant meshes). Our approach is a two steps procedure that defines the desired orientation of the hexahedra with a frame field, then integrates this frame field to generate a deformed 3D grid inside the object. Hexahedra are generated where the grid is non degenerated and not too distorted, and tetrahedra will fill the remaining volume.

7.2.1. Frame Field Generation

A frame field must define the orientation of a cube (the less deformed hexahedron) everywhere inside the object. This object is very difficult to manipulate because it has to be invariant by rotation of 90 degrees around each of its facet normal vector. In [26] we have designed a fast algorithm that is able to define a smooth frame field constrained to be aligned with the object boundary. We represent frames by spherical harmonics (as introduced in [33]) and greatly improve the state of the art thanks to an expression of the boundary constraints that keep the objective function of the optimisation problem very close to quadratic.

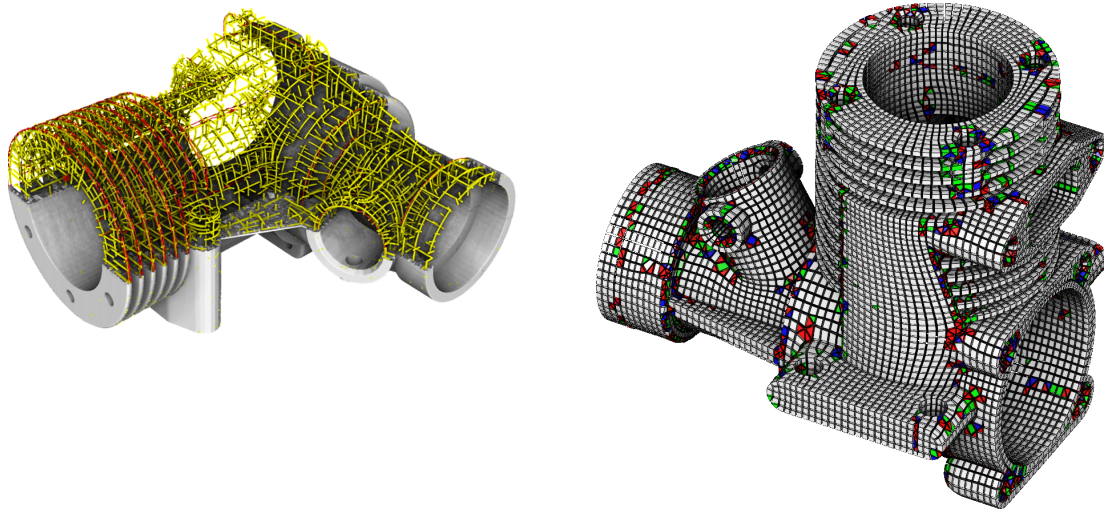


Figure 1. Our Hexahedral dominant meshes generation pipeline have two steps: the generation of a smooth frame field (yellow curves on the left image), and the construction of a mesh with hexahedron that aligns with the frame field.

7.2.2. Generation of Hexahedral-dominant Meshes

The generation of the hex dominant meshes is performed in two steps: place a deformed 3D grid inside the object such that it is aligned with the frame field, then use it to produce hexahedra and fill the rest of the volume with tetrahedra. We developed two solutions for the first step: a 3D extension of PGP [40] and an adaptation of Cubecover [39]. Both solutions have pros and cons so we plan to make them cooperate in a near future. The conversion of this result in an hexahedral-dominant mesh is a very complex problem for which we have a fair solution: we extract a point set from the deformed 3D grid, generate a tetrahedral mesh of the object that is constrained to include the point set in its vertices. From this tetrahedral mesh, we merge sets of tetrahedra into hexahedra with an extension of [37]. We are now working on an alternative solution that will generate hexahedra directly from the deformed 3D grid, and extract the boundary of the rest of the volume as a 2D mesh. From this mesh, we will try to produce more hexahedra by adapting existing combinatorial methods [27].

7.2.3. Impact on FEM Performance

It is admitted by our scientific community that hexahedral meshes are better than tetrahedral meshes for some FEM simulation. We would like to demonstrate evidence of this belief, including fair comparisons with equal running time and/or result accuracy, but the best function basis for each case. For hexahedral dominant meshes, we want to determine if the benefit of using hexahedra deserves having specific function bases devoted to properly link tetrahedral and hexahedral elements. We are developing a new function basis, tailored to non-conformal mixed hexahedra-tetrahedra meshes. Using a combination of tri-linear and quadratic hexahedra, it is possible to construct a space of continuous functions even on a non-conformal mesh. We are now proceeding to analyse the properties of this function space, both theoretically and experimentally. This topic is addressed in the (ongoing) Ph.D. thesis of Maxence Reberol.

7.3. Semi-discrete Optimal Transport in 3D

Participant: Bruno Lévy.

This work introduces a practical algorithm to compute the optimal transport map between a piecewise linear density and a sum of Dirac masses in 3D. In this semi-discrete setting, Aurenhammer *et al.* showed that the optimal transport map is determined by the weights of a power diagram [28]. The optimal weights are computed by minimizing a convex objective function with a quasi-Newton method. To evaluate the value and gradient of this objective function, we propose an efficient and robust algorithm that computes at each iteration the intersection between a power diagram and the tetrahedral mesh that defines the measure. Like in the multilevel proposed by Mérigot, we use a hierarchical algorithm, that uses nested point sets to discretize the source measure.

We think this work may lead to interesting discretizations of the physics, that include the conservation laws (conservation of energy, conservation of momentum ...) deep in their definition, as explained by Jean-David Benamou and Yann Brenier in their fluid dynamics formulation of optimal transport [30].

This work was published in the journal *Mathematical Modeling and Analysis* [10].

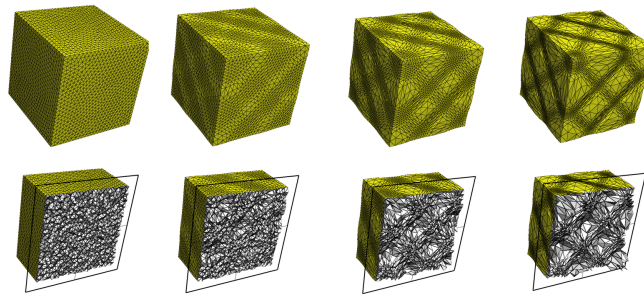


Figure 2. Semi-discrete optimal transport from a constant density to a varying one (product of sines).

7.4. By-example Synthesis of Structurally Sound Shapes

Participants: Jonas Martínez Bayona, Jérémie Dumas, Sylvain Lefebvre.

This is a collaboration with Li-Yi Wei (HKU) on a first project, An Lu (Inria/TU. Munich), Jun Wu (TU. Munich) and Christian Dick (TU. Munich) on a second project.

This work is at the heart of the ERC ShapeForge and considers the by-example synthesis of shapes under structural constraints. We considered two views of the problem that lead to different methodologies.

In a first approach, our goal is to cover a surface with a pattern – an operation akin to texturing in Computer Graphics. The pattern is however used to define the final shape, by determining which parts of the surface are solid or empty. The method operates on a thin voxel shell and does not require any parametrization of the input surface. The pattern is synthesized using a novel formulation for by-example pattern synthesis along surfaces. It is analyzed for structural weaknesses and this information is fed back to the pattern synthesizer, so that seamless reinforcements are added to the structure. We collaborated with researchers from T.U. Munich to analyze the structural behaviour of our structures, and developed a fast evaluation scheme that can be used within our optimization loop to guarantee structural soundness of the resulting design. The work was published in ACM Transactions on Graphics in 2015 [9] (proceedings of SIGGRAPH 2015).

In a second approach we considered the synthesis of shapes that are as rigid as possible under specific boundary conditions and using a prescribed amount of material, while resembling a given input example pattern, as illustrated in Figure 3. Our method is inspired by the field of topology optimization, where rigid shapes are optimized but without any appearance constraints. Our algorithm generates shapes that resemble the input exemplar while being within a user specified percentage of the most rigid shape obtained without the

appearance objective. The work was published in ACM Transactions on Graphics in 2015 [12] (proceedings of SIGGRAPH Asia 2015).



Figure 3. Left. A chair automatically synthesized from a load scenario and an example pattern. The rigidity of the chair is within controlled bounds of a shape optimized without appearance objective. Right. A table design automatically synthesized.

7.5. Modeling for Fabrication

We pursued our research regarding automatic modeling techniques for fabrication, where an algorithm takes into account fabrication constraints to simplify the modeling process. This year we have worked on three projects in this area: the modeling of mechanisms from incomplete 2D definitions, the modeling of self-supporting tight enclosures to assist the fabrication process, and the interactive sculpting of support-free objects.

7.5.1. 3D Fabrication of 2D Mechanisms

Participants: Jean Hergel, Sylvain Lefebvre.

This project considered the automatic modeling of 3D mechanisms from an under-specified 2D model of the mechanism. Our approach casts the synthesis problem as an edge orientation problem in a graph, where graph nodes represent parts of the mechanisms and edges capture their interactions as analyzed by the 2D simulation of the mechanism. The edge orientation determines which parts include which others. Once all inclusions have been determined, we formulate a CSP to solve for the layering problem: each part is assigned a depth 'layer' in 3D. We finally compute the final geometry through CSG (boolean combinations of shapes). This work has been published in Computer Graphics Forum (proceedings of Eurographics 2015) [8]. It received an honorable mention from the best paper committee.

7.5.2. Self-supporting Tight Enclosures

Participants: Samuel Hornus, Sylvain Lefebvre, Frédéric Claux, Jérémie Dumas.

The aim of this project was to develop a technique to automatically generate a tight enclosure in the free space around an object. The challenge was to ensure that the enclosure stays close to the object and be as thin as possible while still being printable without collapsing. Such an enclosure finds at least two important applications : 1. as a protective skin to avoid artifacts when 3D-printing a multi-material object. 2. for generating as-large-as-possible cavities inside the printed object in order to minimize material usage and print time. The work is available as an Inria technical report [22].

7.5.3. Interactive Sculpting of Support-free Objects

Participants: Tim-Christopher Reiner, Sylvain Lefebvre.

Tim Reiner, former PhD student at the Karlsruhe Institute of Technology, joined the team on a Post-Doc position to explore new ideas in the context of modeling, rendering, and fabrication. Starting in March 2015, he developed a voxel-based environment for interactive modeling. In a research project together with Sylvain Lefebvre, our team has derived novel techniques for sculpting support-free 3D shapes. These shapes have the property that they do not require support structures during fabrication on fused deposition modeling or resin-based printers. This work is currently under review.

7.6. Intersection Detection via Gauss Maps; a Review and New Techniques

Participant: Samuel Hornus.

We have revisited the problem of deciding whether two convex objects intersect or not. A systematic view of the problem for polyhedra led us to a unified view of several techniques developed in the computer graphics community and to a new and very fast technique specialized to pairs of tetrahedra. A novel view of the problem as a minimization problem over the 2-sphere led us to the description of new interesting links between the set of planes separating two objects and the silhouette edges of their Minkowski difference. From there, we devised a new algorithm for separating two arbitrary convex objects that is a little bit faster and much more robust than the state of the art technique of Gilbert, Johnson and Keerthi [31]. The work has been summarized in [21].

7.7. Fractal Geometry

Participant: Dmitry Sokolov.

This is a collaboration with Christian Gentil (LE2I), Gilles Gouatay (LSIS), Anton Mishkinis (LE2I).

Additive manufacturing enables for the first time the physical realization of objects having complex geometries. Good approximations of fractals, in particular, can now be manufactured in a variety of materials, including metals. The application domains of fabricated fractal geometries are vast, from the design of “fractal” micro-strip antennas, to the creation of meta-materials.

The main challenge with traditional fractals is the control of the resulting geometry. For example, it is quite challenging to get the desired shape using the system of fractal homeomorphisms proposed by Barnsley [29]. We elaborate here a new type of modeling system, using the facilities of existing CAGD software, while extending their capabilities and their application areas. This new type of modeling system will offer designers (engineers in industry) and creators (visual artists, stylists, designers, architects, etc.) new opportunities to design and produce a quick mock-up, a prototype or a single object. Our approach is to expand the possibilities of a standard CAD system by including fractal shapes while preserving ease of use for end users.

This year we published two papers on the subject [20], [16].

ALPAGE Project-Team

7. New Results

7.1. Playing with DyALog-based parsers

Participants: Éric Villemonte de La Clergerie, Nicholas Parslow.

Éric de la Clergerie has continued the development of two DyALog-based parsers, namely DYALOG-SR, a transition-based dependency parser, and FRMG, a wide-coverage French TAG based on an underlying meta-grammar.

The coverage of FRMG has been extended to cover more (rare) syntactic phenomena. A new conversion scheme has been added for the French version of the Universal Dependency Scheme. Preliminary evaluation experiments have been conducted on the French UD corpus, with both FRMG and DYALOG-SR. FRMG has also been evaluated on the French SPMRL corpus, alone and with coupling with DYALOG-sr

A new notion of secondary edges has been investigated in FRMG metagrammar and parser to provide additional dependency edges, helpful for understanding parsing outputs. In particular, secondary edges are used to denote controls between a verb and its hidden subject.

FRMG's disambiguation tuning is learned from CONLL-like treebanks using supervised learning method. We have conducted preliminary experiments to use unsupervised learning methods with observed accuracy gains between 1 to 1.5 points w.r.t. the no tuning case. However, trying to mix supervised and unsupervised methods have shown no significant gain w.r.t. the supervised case.

The hybridation of FRMG and DYALOG-SR have been tried on a larger spectrum of treebanks.

FRMG has also been exploited during the Master internship of Nicholas Parslow about the use of NLP tools to provide feedback information and correlations on essays written by non-native French learners. In particular, the correction mechanism of FRMG has been extended to cover more cases of frequent errors and provide more explicit messages.

7.2. Linear-time discriminant syntactico-semantic parsing

Participants: Benoit Crabbé, Maximin Coavoux, Rachel Bawden.

In this module we study efficient and accurate models of statistical phrase structure parsing. We focus on linear time lexicalized parsing algorithms (shift reduce) with approximations entailing linear time processing. The existing prototype involves a global discriminant parsing model of the large margin family (Perceptron, Mira, SVM) able to parse user defined structured input tokens [62]. Thus the model can take into account various sources of information for taking decisions such as word form, part of speech, morphology or semantic classes inter alia.

Our model has been generalized in a multilingual setting where we are among the state of the art systems and state of the art on some languages [23]. To our knowledge the parser is one of the fastest existing multilingual phrase structure parser. In order to ease model design for multilingual settings, we currently study efficient feature selection procedures for automating model adaptation to new languages.

We have also extended our model to continuous representations by means of deep learning methods. We currently have a neural network based decision procedure for parsing [22]. It involves both greedy search and beam based search techniques. Current work focuses on the design of dynamic oracles for improving greedy search procedures. This framework is currently tested in the multilingual setting too.

Further work involves to tackle the knowledge acquisition bottleneck problem by integrating either symbolic knowledge such as dictionaries or semi-supervised procedures for improving the formal representation of lexical dependencies in order to leverage data sparsity and estimation issues recurrent in lexicalized parsing.

7.3. French Deep Syntactic Dependency Parsing

Participants: Corentin Ribeyre, Djamé Seddah, Éric Villemonte de La Clergerie, Marie Candito.

At Alpage, we used two distinct but complementary approaches to parse and produce deep syntactic dependency graphs from the DeepSequoia and the DeepFTB (crossref here). The first one was developed by using OGRE [87], [86], a graph rewriting system (crossref here). We developed a set of rewriting rules to transform surfacic syntactic dependency trees into deep syntactic dependency graphs, then we applied this set of rules on previously parsed surfacic trees. Those trees were produced using up to three different surfacic syntactic parsers: FRMG [109], DyALog-SR [109] and Mate [47]. The results were convincing and on par with what we got on English.

The second approach was based on the work made last year regarding the English broad-coverage semantic dependency parsing. We reused our two graph parsers (the first one is based on a previous work on DAG parsing [89] and the second one on the FRMG surfacic syntactic parser [109]) to parse the same graphs. As we previously have shown on English, the use of a mix of syntactic features (tree fragments from a constituent syntactic parser [80], dependencies from a syntactic parser [47], elementary spinal trees using a spine grammar [102], etc.) improve our results. Our intuition is that syntax and semantic are not independent of each other and using syntax could improve semantic parsing. Finally, we extended a dual-decomposition third-order graph parser [76] to incorporate our syntactic feature set and we were able to reach the best performances to this day on the task for both English [28] and French (Ribeyre et al, to appear).

7.4. Towards a French FrameNet

Participants: Marie Candito, Marianne Djemaa, Benoît Sagot.

The ASFALDA project ⁰ is an ANR project coordinated by Marie Candito. 5 partners collaborate on the project, on top of Alpage : the Laboratoire d'Informatique Fondamentale de Marseille(LIF), the Laboratoire de Linguistique Formelle (LLF), the MELODI team (IRIT - Toulouse) and the CEA-List. The project started in October 2012, and will end in march 2016. Its objective is to build semantic resources (generalizations over predicates and over the semantic arguments of predicates) and a corresponding semantic analyzer for French. We chose to build on the work resulting from the FrameNet project [45], ⁰ which provides a structured set of prototypical situations, called *frames*, along with a semantic characterization of the participants of these situations (called *frame elements*). The resulting resources will consist of :

1. a French lexicon in which lexical units are associated to FrameNet frames,
2. a semantic annotation layer added on top of existing syntactic French treebanks
3. and a frame-based semantic analyzer, focused on joint models for syntactic and semantic analysis.

In 2015, we continued the corpus annotation phase, which started in 2014. We currently have about 90 frames and 790 lexical units with at least one annotated occurrence, totalizing about 11, 000 annotated occurrences. We also set up :

- procedures for checking the coherence of the annotations
- a procedure for extracting the "annotated lexicon", namely extract quantitative information about the annotated lexical units, and syntax/semantics interface information (in terms of the probabilistic distributions of the syntactic paths used to express a given semantic role)
- the graphical vizualization of the annotated corpus

We also just started a collaboration with the LIF laboratory for using deep syntactic representations for predicting semantic frames and roles.

7.5. Development of Verb \ni net

Participants: Laurence Danlos, Quentin Pradet, Lucie Barque.

⁰<https://sites.google.com/site/anrasfalda/>

⁰<https://framenet.icsi.berkeley.edu/>

VerbNet is an English lexical resources for verbs, which is internationally known and widely used in numerous NLP applications [74]. Verb \ni net is a French adaptation of this resource. It is semi-automatically developed thanks to the use of two French existing resources created in the 70's: LG, Lexique-Grammaire developed at LADL under the supervision of Maurice Gross, and LVF, Lexique des verbes du français by Dubois and Dubois-Charlier. The idea is to map English classes, which gather verbs with a common syntactic and semantic behavior, into classes of LG and LVF, then to manually adapt the syntactic frames according to French grammar while keeping the thematic roles and the semantic information, [84], [68] [14]. A first version of this work has been achieved in June 2015 in collaboration with Takuya Nakamura (Institut Gaspard Monge) [33].

The next step was to verify the coherence of the resource. A particular focus has been to check the way alternations have been encoded and to document this encoding. A journal article extracted from this documentation has been submitted to the *TAL* journal and Verb \ni net will be released after getting the feedback of the editorial board.

7.6. Development of the French Discourse TreeBank (FDTB)

Participants: Laurence Danlos, Margot Colinet, Jacques Steinlin, Pierre Magistry.

FDTB1 is the first step towards the creation of the French Discourse Tree Bank (FDTB) with a discourse layer on top of the syntactic one which is available in the French Tree Bank (FTB). In this first step, we have identified all the words or phrases in the corpus that are used as “discourse connectives”. The methodology was the following: first, we highlighted all the items in the corpus that are recorded in LexConn [88], a lexicon of French connectives with 350 items, next we eliminated some of these items with the following criteria:

1. first, we filtered out the LexConn items that are annotated in FTB with parts of speech incompatible with a connective use, e.g. *bref* annotated as *Adj* instead of *Adv*, *en fait* annotated as *Pro V* instead of (compound) *Adv*;
2. second, as we lay down for theoretical and practical reasons that elementary arguments of connectives must be clauses or VPs, we filtered out e.g. LexConn prepositions that introduce NPs;
3. last, we filtered out LexConn prepositions and adverbials with a non-discursive function.

The last criterion requires a manual work contrarily to the two others. For example the preposition *pour* (*to*), is ambiguous between a connective use (*Fred s'est dépêché pour être à la gare à 17h* (*Fred hurried to be at the station at 17h*)) and a preposition introducing a complement (*Fred s'est dépêché pour aller à la gare* (*Fred hurried to go to the station*)), and the disambiguation between the two uses is subtle and so the topic of a long paper [58], whose results have been used to enhance Lefff [93].

FDTB1 identifies 9 833 discourse connectives (among 18 535 sentences). This resource is freely available and has been released in May 2015 [36].

FDTB2 is the next step in the creation of the FDTB. It consists in annotating the arguments of the discourse connectives identified in FDTB1 as well as the senses of these connectives (senses expressed through a set of discourse relations). This resource is still worked on.

7.7. Discourse Parsing

Participants: Chloé Braud, Laurence Danlos.

Discourse parsing goal is to reflect the rhetorical structure of a document, how pieces of text are linked in order to form a coherent document. Understanding such links could benefits to several other natural language applications (summarization, language generation, information extraction...).

A discourse parser corresponds to two major subtasks: a segmentation step wherein discourse units (DUs) are extracted, and a parsing step wherein these DUs are (recursively) related through “discourse (rhetorical) relations”. The most difficult task in discourse parsing is the labeling of the relations between DUs, especially when no so-called connective overtly marks the relation (we then talk about implicit relations as opposed to explicit ones).

In her PhD, defended in December 2015, Chloé Braud develops a discourse relation classifier, carrying experiments on French and English. Focusing on the problem on implicit relation identification, this work explores ways of using raw data in combination with the available manually annotated data: this work led to systems based on domain adaptation methods exploiting automatically annotated explicit relations – demonstrating improvements on the French corpus Annodis and on the English corpus PDTB –, and to systems using word embeddings built from raw text to efficiently transform a word based representation of the data – leading to state-of-the art performance or above on the English corpus PDTB without the need of hand-crafted resources [21].

7.8. Towards a morpho-semantic resource for French designed for Word Sense Disambiguation

Participant: Lucie Barque.

The most promising WSD methods are those relying on external knowledge resources [78] but semantic resources for French are scarce. Moreover, existing resources offer fine grained sense distinctions that do not fit to WSD. Our aim is to provide the NLP community with a broad-coverage morpho-semantic lexicon for French that relies on coarse-grained sense distinctions for polysemic units. Preliminary results concern nouns, on which we have first focused because their semantic description, compared to verbs, crucially lacks (for information retrieval, for instance) and because the regular polysemy phenomenon (recurring cases of polysemy within semantic classes) mainly occurs in nominal semantic classes:

- We proposed a linguistically motivated description of general semantic labels for nouns, that will allow for coarse-grained sense distinctions [107]
- Regular polysemy of nouns that can denote an event or a participant of this event has also been described for a large number of French nouns in [46]
- From a morphological point of view, nouns denoting events in French are mostly deverbal nouns (eg. *conversation* 'conversation', *promenade* 'stroll'), but there are also underived event nouns (eg. *guerre* 'war', *séisme* 'earthquake'). We compared their semantic properties in [35].
- Some lexical meanings are not easily captured by ontological semantic classes and a closer look has to be taken at them. Relational meanings in relational nouns are one of them [15].

7.9. Development of the Corpus de Référence du Français

Participants: Stéphane Riou, Benoît Sagot.

The 'Initiative Corpus de Référence du Français' (ICRF) is a project of Institut de Linguistique Française (ILF-FR2393 CNRS), coordinated by its director Franck Neveu and by Benoît Sagot.

The purpose of the ICRF is the development of a first prototype of the future French Reference Corpus, so as to assess the feasibility of this project and evaluate its potential impact. ICRF reuses existing freely-available French corpora, supplemented by additional data in an opportunistic fashion (e.g. a French media critic corpus and the corpus of talks given at an workshop on ethics and neurodegenerative diseases). ICRF preserves copyright and authorship of all corpora used. These corpora have been or will be part-of-speech tagged with MELt, converted to XML (TEI-P5-compliant) and made accessible via a web interface. The aim of ICRF is not to replace individual corpora and the interface will therefore allow, whenever possible, to easily recover access to each individual corpus. ICRF adds 5 metadata tags to categorize each individual corpus: spoken/written, text type and genre, linguistic competence level, date and linguistic area.

In 2015, the normalisation, tagging and conversion to XML of individual corpora has started, following the design of format specifications. The development of the web interface has already started, and a prototype is now available. Users can perform queries (search by tokens and/or POS) and use basic linguistic tools on the corpora (e.g. a concordancer). It is therefore more than a simple search interface or a download site: it improves research and selection of corpus.

7.10. Word order variation in Old French

Participants: Benoit Crabbé, Alexandra Simonenko.

As participant of the strand *Experimental Grammar* of the Labex EFL project *Empirical Foundations of Linguistics*⁰ we study word order issues on Old French and more specifically the relative ordering of complements of ditransitive verbs. The inquiry seeks to identify several factors influencing the ordering of Old French complementation in different texts (varying in dates and genres) by carrying quantitative and statistical work from annotated Old French data.⁰

The first quantitative results [29] will be compared with what is known from corpus studies on the relative ordering of subject and complement in Old French [75]. It will also be compared to the quantitative results obtained on the relative ordering of complements of ditransitive verbs in Modern French [8] and modern English [53]. This comparative perspective is expected to provide new insights on French language evolution.

7.11. Cross linguistic factors governing word order

Participant: Benoit Crabbé.

In many languages, flexible word order often has a pragmatic role and marks the introduction of new information, a focus or a topic shift. Other cases of language-internal word order variation are alternations between two options such as *Mary gave John a book* and *Mary gave a book to John*, which are conditioned on syntactic and semantic factors such as the complexity of the constituents (as in *Mary gave John a book she had read ten times*), their animacy or the meaning of the verb [52].

One of the goals of this module is to investigate the connection between the quantitative aspects of word order variation across languages and the quantitative aspects of word order variation within a language. We study the corresponding patterns in language-internal variation by looking at the syntactically annotated corpora of various languages. Focusing on the variation of the internal word order of the noun-phrase as a case study [25], we explore, in collaboration with Kristina Gulordava (PhD at the University of Geneva, former international visitor at Alpage), to which extent a computational corpus-based analysis can provide new evidence not only for empirical, but also for theoretical linguistic research.

⁰ www.labex-efl.org

⁰SRCMF corpus: <http://srcmf.org/>; MCVF: <http://www.voies.uottawa.ca>

AVIZ Project-Team

7. New Results

7.1. Design Considerations for Composite Physical Visualizations

Participants: Mathieu Le Goc [correspondant], Pierre Dragicevic, Samuel Huron, Jean-Daniel Fekete.

Physical visualization has existed for thousands of years, yet the Information Visualization community is just starting to study it. Many current physical visualizations are monolithic, static, and not interactive. Some of them are made of multiple individual objects that can be rearranged in order to represent a variety of informative configurations. We call them composite physical visualizations. A major benefit of such visualizations is that they support modularity and updatability, but their design space is not well understood.

We show [28] that composite physical visualizations can be classified according to two orthogonal dimensions: i) their level of actuation and ii) their manipulability. Among existing systems, some have a high manipulability but no support for actuation, while others are fully actuated but not manipulable. Only a few systems are combining both qualities and none supports both full manipulability and full actuation. We discuss the tradeoffs between these two dimensions, and identify the opportunities and challenges for future research and design.

7.2. Design Considerations for Enhancing Word-Scale Visualizations with Interaction

Participants: Pascal Goffin, Wesley Willett, Jean-Daniel Fekete, Petra Isenberg.

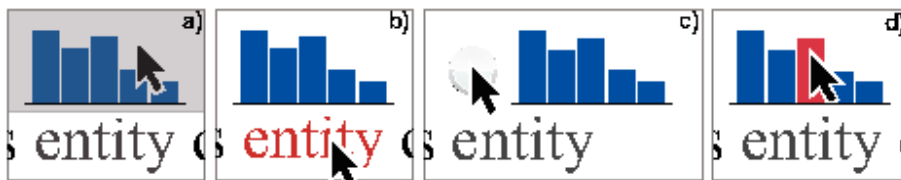


Figure 8. Illustration of where interaction can take place in the context of word-scale visualizations.

This paper presents a design space for interaction with word-scale visualizations. Most sparklines and word-scale visualizations are static and do not support any interaction. However, when word-scale visualizations are used in digital environments, interaction can enhance their use by allowing various data manipulation and management operations. Our design space covers where (Figure 8), when, and how interaction can be triggered for word-scale visualizations embedded in a text document. It also includes how and when to transition from a view where the text with word-scale visualizations is the focus (document-centric view) to a view in which the visualizations becomes the reading focus (visualization-centric view).

7.3. Drawing Characteristics for Reproducing Traditional Hand-Made Stippling

Participants: Domingo Martín, Vicente Del Sol, Celia Romo, Tobias Isenberg [correspondant].

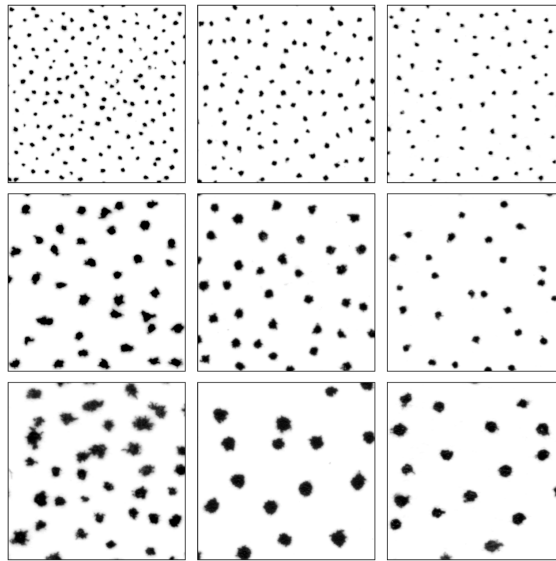


Figure 9. Samples from the stipple dot experiments.

We contribute an in-depth analysis of the characteristics of traditional stippling and relate these to common practices in NPAR stippling techniques as well as to the abilities and limitations of existing printing and display technology. We focus specifically on the properties of stipple dots and consider the dimensions and attributes of pens and paper types used in artistic practice (see Figure 9). With our analysis we work toward an understanding of the requirements for digital stippling, with the ultimate goal to provide tools to artists and illustrators that can replicate the stippling process faithfully in the digital domain. From the results of our study we provide a dataset for use in new example-based stippling techniques, derive a taxonomy of characteristics and conditions for the reproduction of stippling, and define future directions of work.

More on the project Web page: <http://tobias.isenberg.cc/VideosAndDemos/Martin2015DCR>.

7.4. Evaluation of an IEC Framework for Guided Visual Search

Participants: Nadia Boukhelifa [correspondant], Anastasia Bezerianos, Waldo Cancino, Evelynne Lutton.

We evaluated and analysed a framework for Evolutionary Visual Exploration (EVE) [13] (Figure 10) that guides users in exploring large search spaces. EVE uses an interactive evolutionary algorithm to steer the exploration of multidimensional datasets towards two dimensional projections that are interesting to the analyst. This method smoothly combines automatically calculated metrics and user input in order to propose pertinent views to the user. We revisited this framework and a prototype application that was developed as a demonstrator, and summarized our previous study with domain experts and its main findings. We then reported on results from a new user study with a clear predefined task that examined how users leverage the system and how the system evolved to match their needs.

While previously we showed that using EVE, domain experts were able to formulate interesting hypotheses and reach new insights when exploring freely, the new findings indicated that users, guided by the interactive evolutionary algorithm, were able to converge quickly to an interesting view of their data when a clear task was specified. We provided a detailed analysis of how users interact with an evolutionary algorithm and how the system responded to their exploration strategies and evaluation patterns. This line of work aims at building a bridge between the domains of visual analytics and interactive evolution. The benefits are

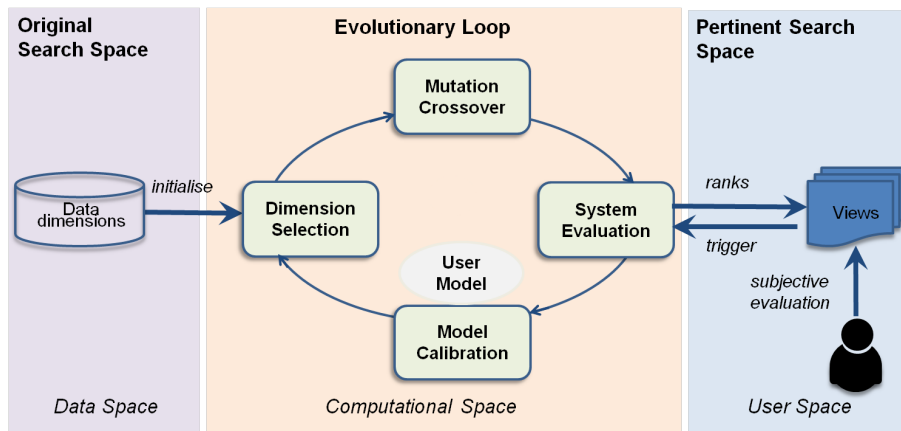


Figure 10. Illustration of the evolutionary visual exploration framework.

numerous, in particular for evaluating Interactive Evolutionary Computation (IEC) techniques based on user study methodologies.

Next, we summarized and reflected upon our experience in evaluating our guided exploratory visualization system [34]. This system guided users in their exploration of multidimensional datasets to pertinent views of their data, where the notion of pertinence is defined by automatic indicators, such as the amount of visual patterns in the view, and subjective user feedback obtained during their interaction with the tool. To evaluate this type of system, we argued for deploying a collection of validation methods that are: user-centered, observing the utility and effectiveness of the system for the end-user; and algorithm-centered, analysing the computational behaviour of the system. We reported on observations and lessons learnt from working with expert users both for the design and the evaluation of our system.

More on the project Web page: <http://www.aviz.fr/EVE><http://www.aviz.fr/EVE>.

7.5. Exploring the Effect of Word-Scale Visualizations on Reading Behavior

Participants: Pascal Goffin, Wesley Willett, Anastasia Bezerianos, Petra Isenberg.

We studied how the integration of small visualizations (word-scale visualizations) into a sentence affects reading speed and memorization during a brief reading task. In particular, we were interested in how different placement types—with their inherent text appearance and layout changes—affect readers. We designed a quantitative study in which we gave sentences with or without small visualizations for participants to read (study conditions are shown in Figure 11). Then, we invited them to answer questions on the sentences. We found that the information encoded in the visualizations is more prominent and easily remembered than information in the written text, but that different placement options had little to no effect on reading performance, even if participants had different preferences.

7.6. Exploration of the Brain's White Matter Structure through Visual Abstraction and Multi-Scale Local Fiber Tract Contraction

Participants: Maarten H. Everts, Eric Begue, Henk Bekker, Jos B. T. M. Roerdink, Tobias Isenberg [correspondant].

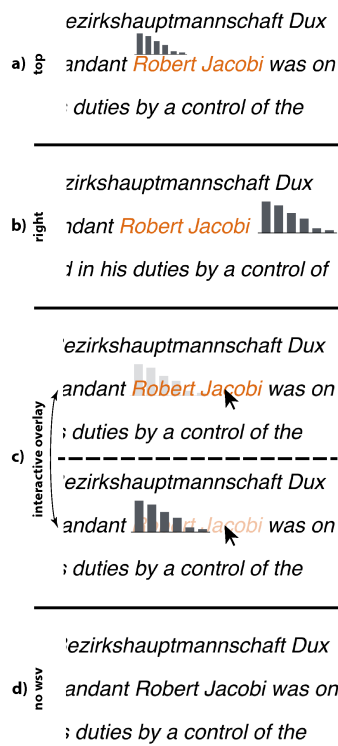


Figure 11. Illustration of the study conditions.

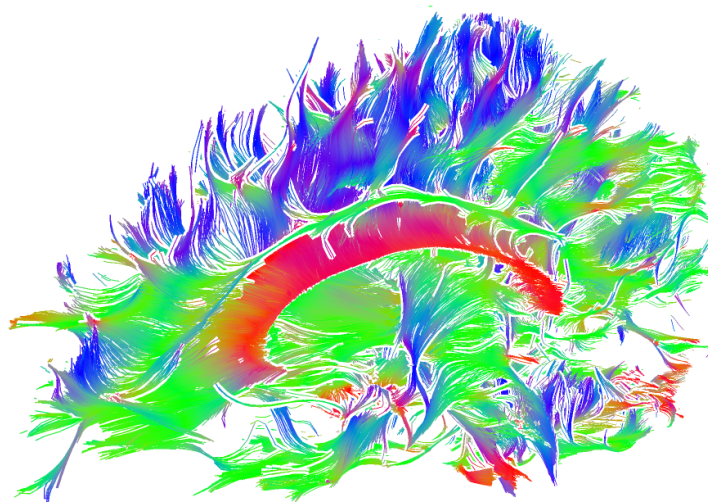


Figure 12. Contraction of the white matter fiber tracts in the brain.

We developed a visualization technique for brain fiber tracts from DTI data that provides insight into the structure of white matter through visual abstraction. We achieve this abstraction by analyzing the local similarity of tract segment directions at different scales using a stepwise increase of the search range. Next, locally similar tract segments are moved toward each other in an iterative process, resulting in a local contraction of tracts perpendicular to the local tract direction at a given scale. This not only leads to the abstraction of the global structure of the white matter as represented by the tracts, but also creates volumetric voids (see Figure 12). This increase of empty space decreases the mutual occlusion of tracts and, consequently, results in a better understanding of the brain's three-dimensional fiber tract structure. Our implementation supports an interactive and continuous transition between the original and the abstracted representations via various scale levels of similarity. We also support the selection of groups of tracts, which are highlighted and rendered with the abstracted visualization as context.

More on the project Web page: <http://tobias.isenberg.cc/VideosAndDemos/Everts2015EBW>.

7.7. Interactive Illustrative Line Styles and Line Style Transfer Functions for Flow Visualization

Participants: Maarten H. Everts, Henk Bekker, Jos B. T. M. Roerdink, Tobias Isenberg [correspondant].

We present a flexible illustrative line style model for the visualization of streamline data. Our model partitions view-oriented line strips into parallel bands whose basic visual properties can be controlled independently. We thus extend previous line stylization techniques specifically for visualization purposes by allowing the parametrization of these bands based on the local line data attributes. Moreover, our approach supports emphasis and abstraction by introducing line style transfer functions that map local line attribute values to complete line styles. With a flexible GPU implementation of this line style model we enable the interactive exploration of visual representations of streamlines. We demonstrate the effectiveness of our model by applying it to 3D flow field datasets (see Figure 13).

More on the project Web page: <http://tobias.isenberg.cc/VideosAndDemos/Everts2015IIL>.



Figure 13. Illustrative line styles applied to streamline visualization of a complex 3D flow.

7.8. Research Agenda for Data Physicalization

Participants: Yvonne Jansen, Pierre Dragicevic [correspondant], Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, Kasper Hornbæk.

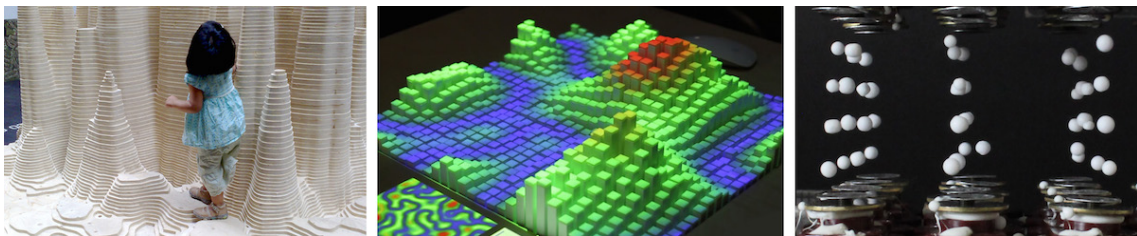


Figure 14. Three examples of data physicalizations.

Physical representations of data have existed for thousands of years. Yet it is now that advances in digital fabrication, actuated tangible interfaces, and shape-changing displays are spurring an emerging area of research that we call Data Physicalization. It aims to help people explore, understand, and communicate data using computer-supported physical data representations. We call these representations physicalizations, analogously to visualizations – their purely visual counterpart. We joined our efforts with research teams from Europe and published a research agenda where we go beyond the focused research questions addressed so far by delineating the research area, synthesizing its open challenges and laying out opportunities for future work. Examples can be seen in Figure 14 .

More on the Data Physicalization Wiki: dataphys.org/.

7.9. Storytelling and Engagement

Participants: Jeremy Boy, Jean-Daniel Fekete, Françoise Detienne.

We conducted three web-based field experiments, in which we evaluated the impact of using initial narrative visualization techniques and storytelling on user-engagement with exploratory information visualizations. We conducted these experiments on a popular news and opinion outlet, and on a popular visualization gallery website. While data journalism exposes visualizations to a large public, we do not know how effectively this public makes sense of interactive graphics, and in particular if people explore them to gain additional insight to that provided by the journalists. In contrast to our hypotheses, our results indicated that augmenting exploratory visualizations with introductory ‘stories’ does not seem to increase user-engagement in exploration.

Many online data graphics use narrative design elements to explain a given dataset in a straightforward and compelling way. According to New York Times graphic editors Mike Bostock and Shan Carter, these explanatory graphics are preferable for data-journalism, as they have the advantage of exposing up-front what the main insights from the data are, without making people ‘have to work for them.’ However, most only provide limited interactivity, which reduces the potential for personal extraction of insight. In essence and by definition, Information visualization (Infovis) is interactive and exploratory. Thus, finding ways to make exploratory graphics more accessible and engaging to people is important, because if open/public/civic data is to truly empower people, then these people should be able to use appropriate tools to gain their own insights and knowledge—not only that provided by journalists in articles written or designed from a specific perspective. We explored the potential of narrative visualization techniques and storytelling to trigger this desired user-engagement. By engagement, we specifically mean a user’s investment in the exploration of a visualization.

AYIN Team

7. New Results

7.1. Markov Random Fields

7.1.1. *New hierarchical joint classification method of SAR and optical multiresolution remote sensing datas*

Participants: Ihsen Hedhli, Josiane Zerubia [contact].

This work was carried out in collaboration with Prof. Gabriele Moser and Prof. Sebastiano Serpico from DITEN departement [www.diten.unige.it/], University of Genoa, Italy.

Nowadays, a wide variety of remote sensing images is available. Therefore, it becomes more and more important to be able to analyze compound data sets consisting of different types of images acquired by different sensors, as they allow a spatially distributed and temporally repetitive view of the monitored area at the desired spatial scales. In particular, the opportunity of joint availability of synthetic aperture radar (SAR) and optical images offers high resolution (HR), all-weather, day/night, short revisit time data, as well as polarimetric and multifrequency acquisition capabilities. Similarly, the strong differences in terms of wavelength range (microwave vs. visible and near infrared), sensitivity to cloud cover and sun illumination (strong for optical imagery vs. almost negligible for SAR), and noise-like properties (speckle in SAR vs. generally low noise variance in current HR optical sensors) make the joint use of HR optical and SAR imagery especially interesting for many applications to environmental monitoring and natural risk management. Within this framework, there is a definite need for classification methods that automatically correlate different sets of images taken on the same area from different sensors and at different resolutions. This year we developed a novel classification approach for multiresolution, multisensor (optical and synthetic aperture radar), and/or multiband images. Accurate and time-efficient classification methods are particularly important tools to support rapid and reliable assessment of the ground changes. Given the huge amount and variety of data available currently from last-generation satellite missions, the main difficulty is to develop a classifier that can benefit from multiband, multiresolution, and multisensor input imagery. As shown in Figure 1, the proposed method addresses the problem of multisensor fusion of SAR with optical data for classification purposes, and allows input data collected at multiple resolutions and additional multiscale features derived through wavelets to be fused. The proposed approach formalizes a supervised Bayesian classifier within a multiple quadtree topology that combines a class-conditional statistical model for pixel-wise information and a hierarchical Markov random field (MRF) for multisensor and multiresolution contextual information.

7.2. Marked point processes

7.2.1. *Integrating RJMCMC and Kalman filters for multiple object tracking*

Participants: Paula Craciun, Josiane Zerubia [contact].

This work has been done in collaboration with Dr. Mathias Ortner from Airbus D&S [<http://www.space-airbusds.com/fr/>]

Recently, we have proposed a new spatio-temporal marked point process model for tracking small, rigid objects in high resolution images. We have shown very good detection and tracking results for synthetic biological data as well as remotely sensed sequences. The model is based on defining a dedicated energy function that is highly non-convex. The solution is found by minimizing this energy function using a suitable batch-optimization scheme based on Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampler. This approach is motivated by the low temporal frequency of the sequences (1Hz).

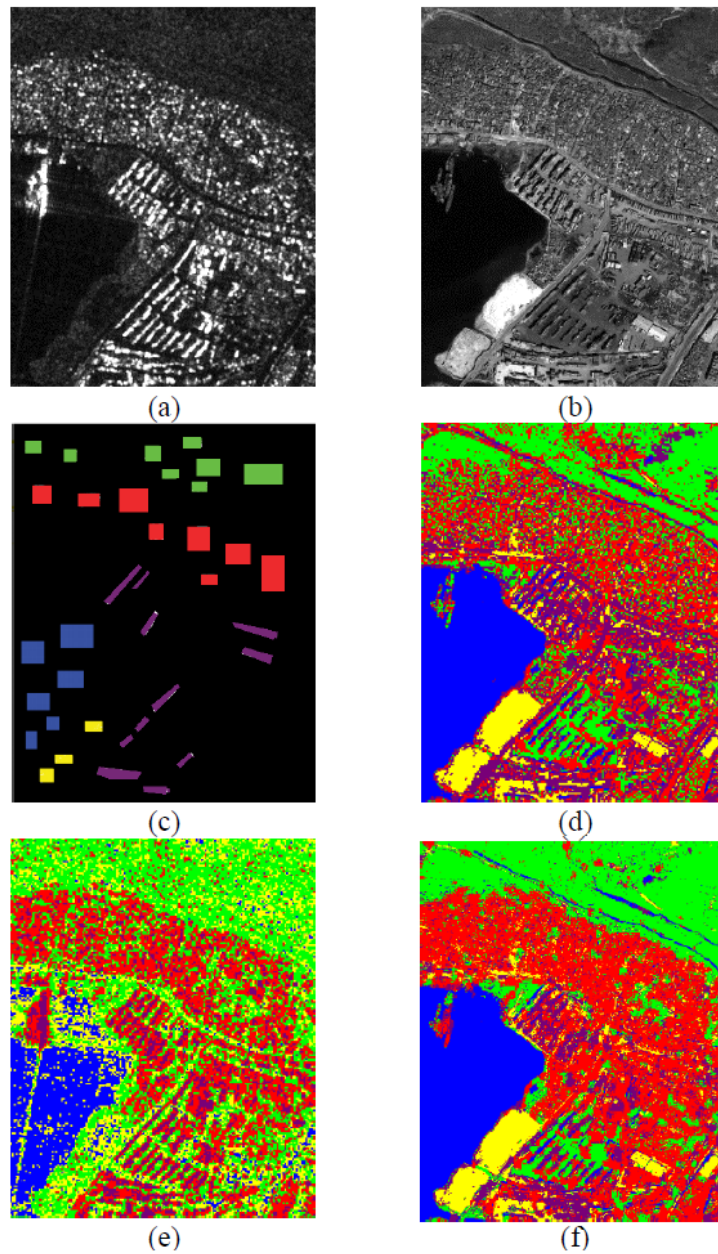


Figure 1. (a) SAR image (© ASI), (b) one channel from the optical image (© GeoEye), (c) the available ground truth, (d) hierarchical MRF-based classification obtained from the optical image, using Laferté method, (e) hierarchical MRF-based classification obtained for the SAR image, using Laferté method, (f) hierarchical MRF-based classification obtained by the proposed method.

Sequential filters have proven to provide relatively fast and reliable tracking performances in particular for single target tracking. We have efficiently exploited the properties of sequential filters within the RJMCMC sampling scheme. The filter is used to generate more meaningful perturbation proposals which are then evaluated using an appropriate Green acceptance ratio. Better perturbation proposals increase the acceptance probability of the overall RJMCMC sampling scheme which in turn leads to a faster convergence.

Figure 2 shows the detection and tracking results on two synthetic biological sequences as well as on two sequences of simulated satellite images of Toulon by courtesy of Airbus Defence & Space, France. The evolution of the energy with the number of iterations for the standard RJMCMC sampler and the proposed sampler is also shown. The proposed sampler is depicted blue.

7.2.2. Initialization and estimation of parameters for marked point processes applied to automatic object detection on satellite images

Participants: Aurélie Boisbunon, Josiane Zerubia [contact].

This work has been done in collaboration with Dr. Rémi Flamary (Université de Nice Sophia Antipolis), Prof. Alain Rakotomamonjy (Université de Rouen) et Alain Giros (CNES). It was partially funded by the French Spatial Agency CNES [<http://www.cnes.fr>].

Sparse representations, large scale, stochastic algorithms, machine learning, image processing Marked point processes (MPP) strongly rely on parameters, whose estimation affects both computation time and performances. In this work, we proposed two approaches: the first one consists in initializing MPPs with a first coarse solution obtained very quickly from sparse regularization methods, while the second one estimates the parameters by the Stochastic Approximation Expectation-Maximization (SAEM) algorithm. We give details on both approaches below. The first coarse solution is obtained from a deterministic sparse regularization method. This method is based on the representation of an image with objects as a sum of convolutions between atoms of a dictionary and matrices of positions of the objects in the image. The atoms of the dictionary are fixed in advance and correspond to different instances of the objects (scales, angles, shapes, etc). This way, we transform the problem of object detection into the problem of estimating extremely sparse matrices. The algorithm we derived for solving the associated optimization problem is both parallelized and very efficient. This work started last year, and continued this year by conducting more tests.

7.2.3. Generic curvilinear structure modeling via marked point process theory

Participants: Seong-Gyun Jeong, Yuliya Tarabalka, Josiane Zerubia [contact].

This work has been done in collaboration with Dr. Nicolas Nisse (COATI team [<https://team.inria.fr/coati/>], Inria-SAM) and Dr. Yuliya Tarabalka (Titane team [<https://team.inria.fr/titane/team/>], Inria-SAM)

We propose a novel curvilinear structure reconstruction algorithm based on ranking learning and graph theory. In this work we reconstruct the curvilinear structure as a set of small line segments (via MPP). Specifically, we infer the structured output ranking of the line segments via Structured Support Vector Machine (SSVM). To predict the existence of the curvilinear structure, we measure oriented image gradient maps and morphological profiles. We propose an orientation-aware curvilinear feature descriptor and a feature grouping operator to improve the structural consistency for learning system. In order to provide topological information, we develop a graph-based curvilinear structure reconstruction algorithm. The proposed algorithm builds a graph based on the output ranking scores and searches the longest geodesic paths which are associated with the latent curvilinear structure. Experimental results (see Figure 3) show that the proposed algorithm faithfully detects the curvilinear structures and preserves topological information compared with the competing algorithms.

7.3. Other approaches

7.3.1. Acne detection on polarized or non-polarized images

Participants: Zhao Liu, Josiane Zerubia [contact].

This work is in collaboration with Dr. Queille-Roussel and Prof. Bahadoran in CHU Nice, France. Now Dr. Zhao Liu is a post-doc at Manchester University [www.manchester.ac.uk/], Manchester, UK.

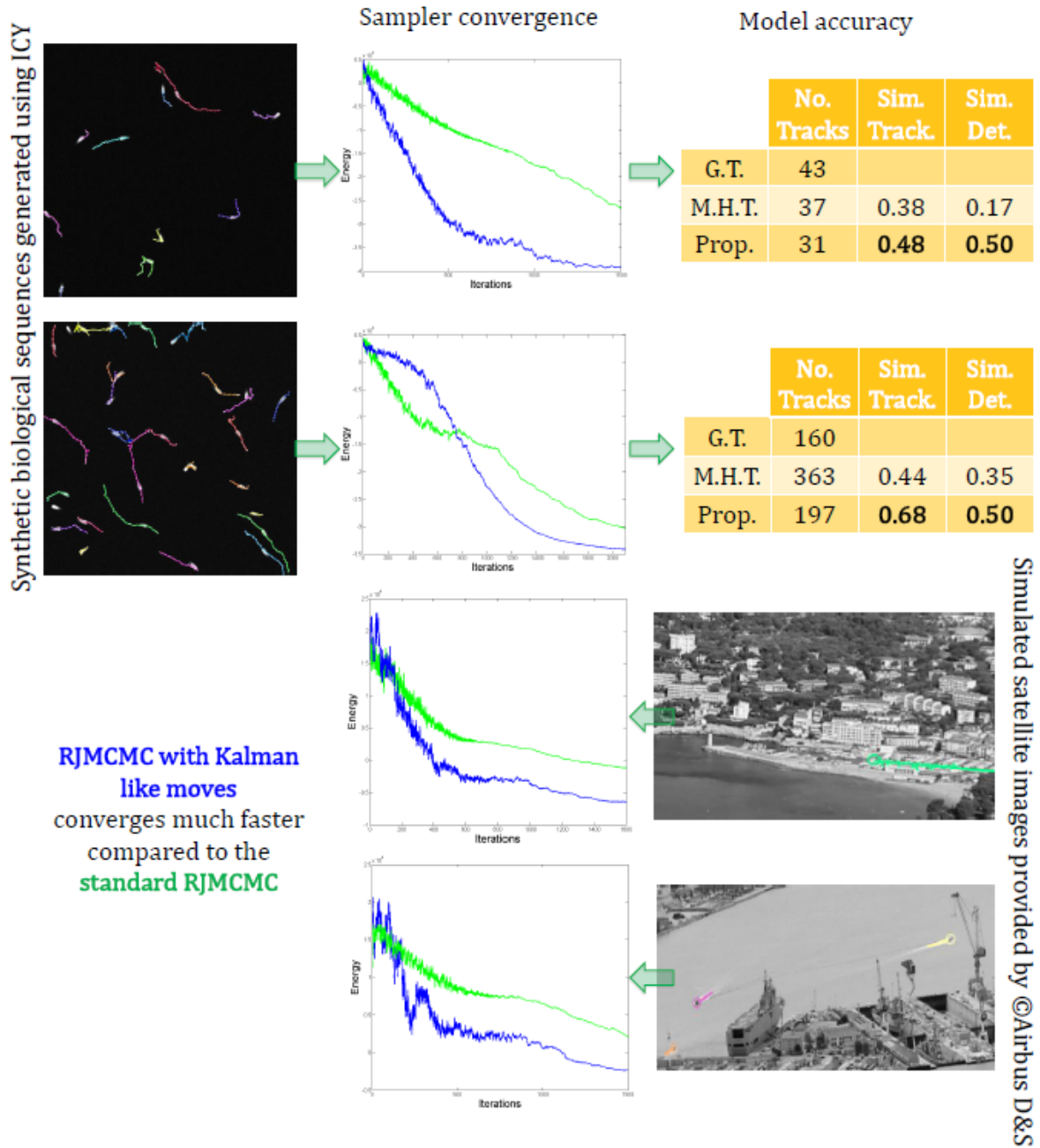


Figure 2. Tracking results and sampler convergence on two synthetic biological sequences (generated using ICY [<http://icy.bioimageanalysis.org/>], a free software offered by the Quantitative Analysis Unit from the Pasteur Institute, France) as well as two sequences of simulated satellite images of Toulon (by courtesy of Airbus Defence & Space, France). The RJMCMC sampler with Kalman like moves (shown in blue) requires a significantly lower number of iterations until convergence as compared to the standard RJMCMC.

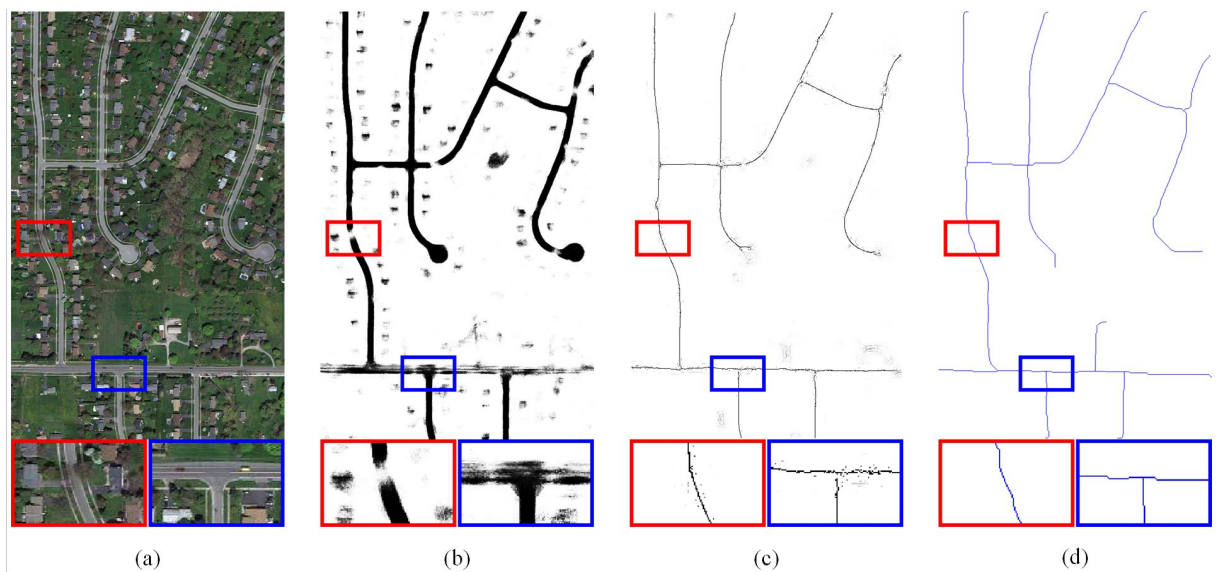


Figure 3. Compared with (b) the segmentation and (c) the centerline detection methods, (d) the proposed algorithm well represents topological features of the curvilinear structure. Setting a threshold value yields to lose correlated information of the pixels on the reconstructed curvilinear structure. In this example, road network is partially occluded by trees or cars, so that the local measure often fails to detect the underlying curvilinear structures. Although the centerline is able to quantify scale (width) of curvilinear structure, it is inaccurate to classify pixels around junctions. In this work we learn spatial patterns of the curvilinear structures with structured output ranking scores. We also propose a graph-based representation algorithm to obtain the topological information.

This work is in collaboration with Dr. Queille-Roussel and Prof. Bahadoran in CHU Nice, France.

Acne vulgaris, a highly prevalent skin disease, has a significant life quality impact on patients. It is generally believed that this type of skin disorder results from proliferation of propionibacterium acnes in pilosebaceous units, which can lead to inflammatory lesions due to increase of oxyhemoglobin level. So far there is no golden standard for acne diagnosis in clinics. It entirely depends on dermatologists' experience for acne assessment. But significant variability among individual diagnosis may lead to less trustworthy results, and less reproducibility of human evaluation makes the comparison of acne follow-up difficult. This work, incorporating the knowledge of optical characteristics of human skin, identifies cutaneous chromophore distribution using bilateral decomposition. Then the inflammatory acne lesions are detected by a Markov random field (MRF) model associating the chromophore descriptors. Experimental results (see Figure 4) show that the proposed method is robust to large dynamic range intensity, and the derived automatic segmentation of inflammatory acne appears to be highly consistent to human visual assessment. This research work was started in 2013. This year, more tests have been conducted on polarized and non-polarized images.



Figure 4. Acne detection using proposed method. (Left) Original image provided by CHU Nice. (Right) Acne detection result.

7.3.2. Finer registration of facial wrinkles in time series images

Participants: Nazre Batool, Josiane Zerubia [contact].

Dr. Batool was funded by the Inria-DPEI fellowship for the period Feb. 2014 – May 2015. Currently she is a postdoc researcher at CMIV, Linköping University [www.liu.se/cmiv], Linköping, Sweden.

The goal of this work is to evaluate quantitatively the subtle variations in facial wrinkles for the same subject in response to treatment using image-based analysis. Any image analysis technique for the analysis of such subtle image variations would require high accuracy and precision for good performance. As in other imaging problems geared towards detection of temporal changes, accurate registration of key image features (wrinkles) is mandatory as a first step. We propose to compare image features in key wrinkle sites only while excluding the noise introduced by changes in surrounding skin texture. Therefore, previously we proposed a 2-step registration algorithm where the initial registration was based on the alignment of facial landmarks such as

corners of eyes, nose, and mouth. Then a method based on Large Deformation Diffeomorphic Metric Mapping (LDDMM) was used to achieve finer local registration for wrinkles. However, the LDDMM algorithm had the shortcoming of the unavailability of time invariant finer facial landmarks and that the deformations were guided by image intensities which were varying among images as well due to subtle changes in skin texture. The deformation of skin due to underlying movement can be categorized loosely as locally rigid because the local skin texture remains constant but globally non-rigid because of the movement of skin areas due to slight expression and misalignment. Due to this dual nature of deformation, registration schemes such as thin plate spline or affine transformations are not applicable. Our improved approach is to guide the LDDMM registration on skin features with higher intensity gradients only (such as due to moles, wrinkles, rough surface) which have the higher probability of being constant and detected across temporal changes. First we detect key landmarks and landmark correspondences using the Gabor feature images where the phase correlation is used to find estimates of landmark correspondences. The phase correlation is based on the well-known Fourier shift property i.e. a shift in the spatial domain of two images results in a linear phase difference in the frequency domain of their respective Fourier Transforms. Figure 5 shows Gabor features of two images captured 4 weeks apart in (a) and (b). Figure 5 (c) shows key landmarks placed at high Gabor amplitude sites and (d) shows their corresponding landmarks detected using Fourier phase correlation.

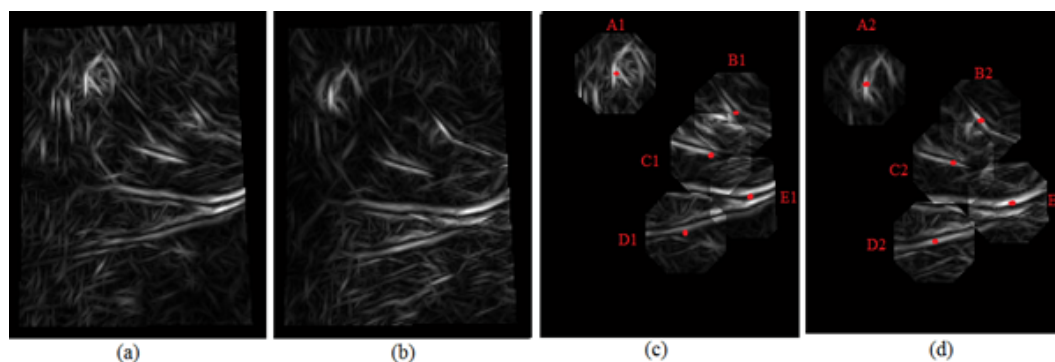


Figure 5. Detection of landmarks and correspondences. (a) Gabor response for source image. (b) Gabor response for target image. (c) 5 landmarks with circular templates in source image. (d) Corresponding detected landmarks with circular templates in target image.

Then, as a next step, the detected key landmarks and their corresponding positions are used in the landmark based LDDMM algorithm to find locally non-rigid deformations between two images. Figure 6 shows an example where the corresponding landmarks are shown as black dots in (a) and (b). Fig. 5 (c) shows the image in (a) wrapped to (b) using LDDMM based on landmark correspondences. In (d) the drifts of landmarks are shown during the LDDMM algorithm and (e) shows the non-rigid deformation of underlying image grid. In the future, the proposed wrinkle registration algorithm will be used to compare wrinkle intensities in time series of images to quantify very minute changes in wrinkles in response to dermatological treatments.

7.3.3. Hyperspectral Image Processing for Detection and Grading of Skin Erythema

Participants: Ali Madooei, Josiane Zerubia [contact].

Ali Madooei worked at Inria Sophia Antipolis on an internship funded by the Canadian Mitacs Globalink Research Award & Inria. He is currently in his last year of PhD at Simon Fraser [www.sfu.ca/] University, Canada. This work has been conducted in collaboration with Ramy M. Abdlaty, Lilian Doerwald-Munoz, Dr. Joseph Hayward and Prof. Qiyin Fang from Mc Master university [http://future.mcmaster.ca/]/Juravinsky cancer center [www.jcc.hhsc.ca/], Canada, and Prof. Joseph Hayward from Simon Fraser University, Canada.

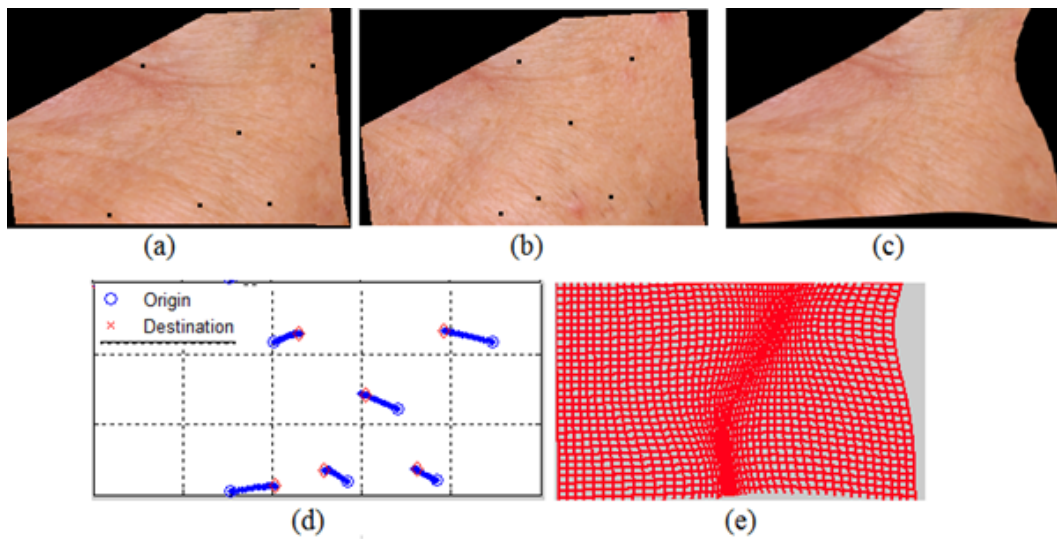


Figure 6. Finer registration of micro features in skin images. (a) Source image with landmarks shown as black dots. (b) Target image with landmarks. (c) Source image wrapped to target image. (d) Path of landmarks during LDDMM registration. (e) Deformation of underlying grid during LDDMM registration.

Acute skin erythema is a common side effect in patients undergoing radiotherapy treatment. It displays itself as an increase in skin redness and irritation. Erythema has been reported to correlate to individual patient response to radiation and therefore may be useful to guide and modify courses of treatment in a timely manner. Currently, upon visual examination, a qualitative score can be assigned to characterize the severity of erythema, which then may be used for assessing radiation response. Due to the subjective nature of this method, additional non-invasive techniques are needed for more accurate evaluation. Previous studies have mainly focused on tissue reflectance spectroscopy or imaging photography. The former retrieves spectral information from point measurements while the latter is obtained with conventional Red, Green, Blue (RGB) colour cameras. Photography has the advantage of offering spatial information but this comes at the cost of losing much of spectral information. We use hyperspectral imaging (HSI) which provides both spatial and spectral representation of the affected area. A hyperspectral camera effectively divides the spectrum into very many thin image slices (the actual number depending on the camera and application see Fig. 7). This fine-grained slicing reveals spectral structure that may not be evident to the eye or to an RGB camera but can provide a rich set of information for image processing. As an emerging imaging modality for medical applications, the combination of HSI devices with adequate image processing techniques offers the perfect landscape for developing new methods for noninvasive disease monitoring and diagnosis.

The purpose of our study was to investigate the possibility of monitoring the degree of erythema using HSI data. To this aim, we proposed an image processing pipeline and conducted controlled experiments to demonstrate the efficacy of the proposed approach for (1) reproducing clinical assessments, and (2) outperforming RGB imaging data. We combined the problem of erythema detection and grading into a multi-class classification problem where each pixel is classified as one of the four erythema classes or a non-erythema class. We used a weighted LDA (linear discriminant analysis) classifier to deal with noisy labels. Moreover, we devised pre-processing steps to deal with noisy measurements. We evaluated the system against the clinical assessment of an experienced clinician. We also compared the performance to that of using digital photography (instead of HSI). The results from this preliminary study are encouraging and indicate that hyperspectral image data do contain relevant information, and indeed outperform imaging photography. In the future, we want to

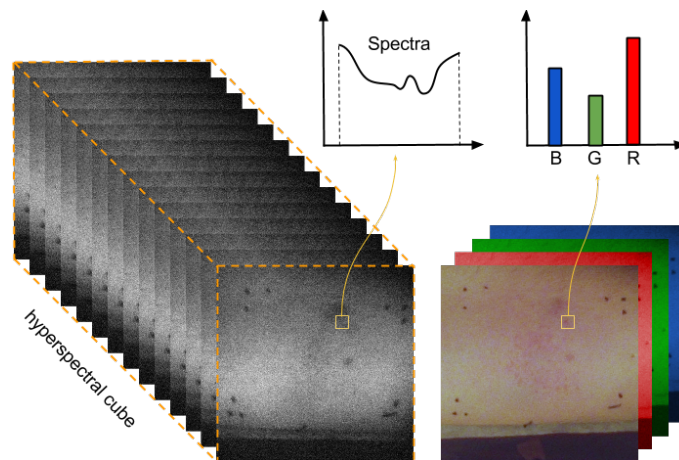


Figure 7. A schematic representation of hyperspectral vs. RGB image data.

extend the technique to further detect other skin responses to radiation (such as dry/moist desquamation, skin necrosis, etc.) and also to experiment with real patients undergoing radiotherapy. Our ultimate objective is to build a system for monitoring radiation response in individuals using HSI technology and image processing.

7.3.4. SAR data classification using generalized Gamma mixture model

Participants: Vladimir Krylov, Josiane Zerubia [contact].

Vladimir Krylov is a former AYIN post-doc, now post-doc at DITEN department, University of Genoa [www.diten.unige.it/], Italy. This work has been performed in collaboration with Prof. Heng-Chao Li, Prof. Ping-Zhi Fan (Southwest Jiaotong University, Chengdu [english.swjtu.edu.cn/], China) and Prof. William Emery (University of Colorado [www.colorado.edu/], Boulder, USA).

The accurate statistical modeling of synthetic aperture radar (SAR) images is a crucial problem in the context of effective SAR image processing, interpretation and application. In this work a semi-parametric approach is designed within the framework of finite mixture models based on the generalized Gamma distribution (G Γ D) in view of its flexibility and compact analytical form. Specifically, we have developed a generalized Gamma mixture model (G Γ MM) to implement an effective statistical analysis of high-resolution SAR images and proved the identifiability of such mixtures. A low-complexity unsupervised estimation method has been derived by combining the proposed histogram-based expectation-conditional maximization algorithm and the Figueiredo-Jain mixture estimation algorithm. This resulted in a numerical maximum likelihood (ML) estimator that can simultaneously determine the ML estimates of component parameters and the optimal number of mixture components. The state-of-the-art performance of the proposed method has been validated experimentally on a wide range of high-resolution SAR amplitude and intensity images.

In Fig. 8 we demonstrate a typical result of the developed statistical modeling technique on a portion of a 2 meter resolution L-band image acquired by an airborne EMISAR system. The unsupervised G Γ MM estimate contains five components and reports a very accurate result that outperforms the considered benchmark statistical modeling methods. In order to visualize the estimated five statistical components we also report a maximum likelihood classification map.

7.3.5. Multitemporal image change detection with a False Discovery Rate approach

Participants: Vladimir Krylov, Josiane Zerubia [contact].

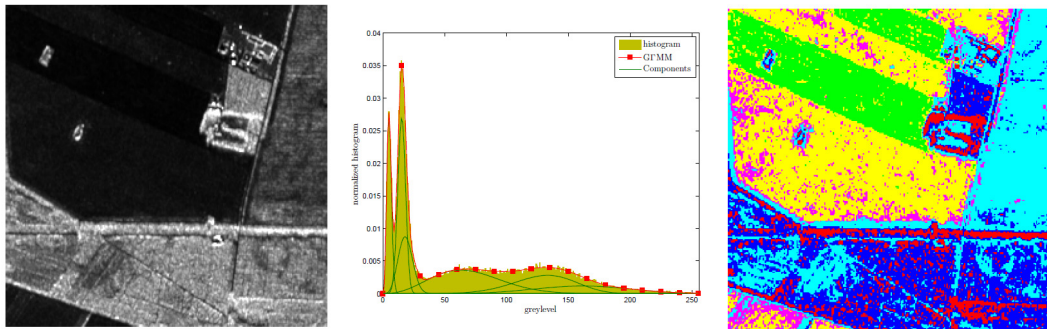


Figure 8. Statistical modeling of a EMISAR (©ESA) image (left) by generalized Gamma mixture model (middle) and its visualization by maximum likelihood classification (right).

This work has been performed in collaboration with Prof. Sebastiano Serpico and Prof. Gabriele Moser, DITEN department, University of Genoa [www.diten.unige.it/], Italy.

Multitemporal change detection is one of the fundamental image processing problems and multiple detection, monitoring and tracking applications rely on its accurate and timely performance. In this work we address the problem of unsupervised change detection on two or more coregistered images of the same object or scene at several time instants. The designed method is appropriate for short image sequences with a relatively small amount of changes. Such analysis is instrumental in various applications where acquisitions are relatively sparse and report limited meaningful changes, in particular, in remote sensing and medical image processing. We develop a novel patch-based hypothesis testing approach which is based on a false discovery rate formulation for statistical significance testing. This alternative error metric allows to adjust the family-wise error rate by imposing control over the proportion of the false positives in the detection. The designed change detector allows the use of various statistical features. The appropriate choice of the latter enables the detector to address application-specific detection problems with a particular set of disturbance factors, like noise, illumination variation, etc. In particular, we demonstrate the use of two rank-based statistics for change detection on image pairs and one multisample statistic for the analysis of image sequences. The experiments with remotely sensed radar, dermatological, and still camera surveillance imagery demonstrate competitive performance and flexibility of the proposed method.



Figure 9. Change detection on a pair of 15 meter resolution XSAR images (first and second) obtained with a false discovery rate error metric based on the Cramer-von Mises statistic. The changes are highlighted with red circles (second), and the unsupervised detection result is reported in black (third).

A typical result obtained with the proposed change detection technique is reported in Fig. 9. The proposed approach gives a unified statistical thresholding procedure to perform change detection based on statistical

features that have a known distribution under the no-change hypothesis. This approach is essentially non-parametric and is highly parallelizable.

Chroma Team

7. New Results

7.1. Sensor Fusion

7.1.1. Observability properties of the visual-inertial structure

Participant: Agostino Martinelli.

We continued to investigate the visual-inertial structure from motion problem by further addressing the following issues:

1. analytically deriving its observability properties in challenging scenarios (i.e., when some of the system inputs are unknown and act as disturbances);
2. obtaining simple and efficient methods for data matching and localization.

Regarding the first issue, we extended our previous results (published last year on the journal *Foundations and Trends in Robotics* [43]) by also including the extreme case of a single point feature and when the camera is not extrinsically calibrated. Even if this extension seems to be simple, the analytic computation must be totally changed. Indeed, by including in the state the camera extrinsic parameters, the computation, as carried out in [43] in the case when the camera is calibrated, becomes prohibitive.

The problem of deriving the observability properties of the visual-inertial structure from motion problem, when the number of inertial sensors is reduced, corresponds to solve a problem that in control theory is known as the Unknown Input Observability (UIO). This problem is still unsolved in the nonlinear case. In [43] we introduced a new method able to provide sufficient conditions for the state observability. On the other hand, this method is based on a state augmentation. Specifically, the new extended state includes the original state together with the unknown inputs and their time-derivatives up to a given order. Then, the method introduced in [43] is based on the computation of a codistribution defined in the augmented space. This makes the computation necessary to derive the observability properties dependent on the dimension of the augmented state and consequently prohibitive in our case. Our effort to deal with this fundamental issue, was devoted to separate the information on the original state from the information on its extension. We fully solved this problem in the case of a single unknown input. For the general case, we partially solved this problem and we suggested a technique able to partially perform this separation. Since these results are very general (their validity is not limited to the visual-inertial structure from motion problem) we presented them at two international conferences on automatic control (SIAM on Control and Applications, [18] and MED, [16]). By applying these new methods to the the visual-inertial structure from motion problem, we obtained the following result. Even in the case of a single point feature, the information provided by a sensor suit composed by a monocular camera and two inertial sensors (along two independent axes and where at least one is an accelerometer) is the same as in the case of a complete inertial measurement unit (i.e., when the inertial sensors consist of three orthogonal accelerometers and three orthogonal gyroscopes). This result has been presented at ICRA, [17].

Regarding the second issue, our focus was in the framework of Micro Aerial Vehicle navigation. State of the art approaches for visual-inertial sensor fusion use filter-based or optimization-based algorithms. Due to the nonlinearity of the system, a poor initialization can have a dramatic impact on the performance of these estimation methods. Last year, we published, on the journal of computer vision, a closed-form solution providing such an initialization [42]. This solution determines the velocity (angular and linear) of a monocular camera in metric units by only using inertial measurements and image features acquired during a short time interval. This year, we study the impact of noisy sensors on the performance of this closed-form solution. Additionally, starting from this solution, we proposed new methods for both localization and data matching in the context of micro aerial navigation. These methods have been tested in collaboration with the vision and perception team in Zurich (in the framework of the ANR-VIMAD) and published on the journal of *Robotics and Autonomous Systems* [4].

7.1.2. Sensing floor for Human & objects localisation and tracking

Participants: Mihai Andries (inria Nancy, Larsen), Olivier Simonin, François Charpillat (inria Nancy, Larsen).

In the context of the PhD of Mihai Andries, co-advised by François Charpillat (Inria Nancy, Larsen) and Olivier Simonin, we investigated a large distributed sensor — a grid of connected sensing tiles on the floor — that was developed by the Maia team, at Nancy, in 2012.

Localization, tracking, and recognition of objects, robots and humans are basic tasks that are of high value in the applications of ambient intelligence. Sensing floors were introduced to address these tasks in a non-intrusive way. To recognize the humans moving on the floor, they are usually first localized, and then a set of gait features are extracted (stride length, cadence, and pressure profile over a footstep). However, recognition generally fails when several people stand or walk together, preventing successful tracking. In the PhD, defended on December 15 [27], we proposed a detection, tracking, and recognition technique which uses objects' weight. It continues working even when tracking individual persons becomes impossible. Inspired by computer vision, this technique processes the floor pressure-image by segmenting the blobs containing objects, tracking them, and recognizing their contents through a mix of inference and combinatorial search. The result lists the probabilities of assignments of known objects to observed blobs. The concept was successfully evaluated in daily life activity scenarii, involving multi-object tracking and recognition on low-resolution sensors, crossing of user trajectories, and weight ambiguity. This model can be used to provide a probabilistic input for multi-modal object tracking and recognition systems. The model and the experimental results have been published in Journal IEEE Sensors [1] and international conference ICRA 2015 [7].

7.2. Bayesian Perception

Participants: Christian Laugier, Lukas Rummelhard, Amaury Nègre, Jean-Alix David, Procópio Silveira-Stein, Jerome Lussereau, Tiana Rakotovaio, Nicolas Turro (sed), Jean-François Cuniberto (sed), Diego Puschini (cea Dacle), Julien Mottin (cea Dacle).

7.2.1. Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT)

Participants: Lukas Rummelhard, Amaury Nègre, Christian Laugier.

In 2015, the research work on *Bayesian Perception* has been done as a continuation and an extension of some previous research results obtained in the scope of the former Inria team-project e-Motion. This work exploits the *Bayesian Occupancy Filter (BOF)* paradigm [28], developed and patented by the team several years ago⁰. It also extends the more recent concept of *Hybrid Sampling BOF (HSBOF)* [46], whose purpose was to adapt the concept to highly dynamic scenes and to analyse the scene through a static-dynamic duality. In this new approach, the static part is represented using an occupancy grid structure, and the dynamic part (motion field) is modeled using moving particles. The *HSBOF* software has been implemented and tested on our experimental platforms (equipped Toyota Lexus and Renault Zoe) in 2014 and 2015; it has also been implemented in 2015 on the experimental autonomous car of Toyota Motor Europe in Brussels.

The objective of the research work performed in 2015 was to overcome some of the shortcomings of the *HSBOF* approach⁰, and to obtain a better understanding of the observed dynamic scenes through the introduction an additional object level into the model. The new framework, whose development will be continued in 2016, is called *Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT)* [10]. This work has mainly been performed in the scope of the project *Perfect* of IRT Nanoelec⁰ (financially supported by the French ANR agency⁰), and also used in the scope of our long-term collaboration with Toyota.

⁰The *Bayesian programming formalism* developed in e-Motion, pioneered (together with the contemporary work of Thrun, Burgards and Fox [53]) a systematic effort to formalize robotics problems under Probability theory—an approach that is now pervasive in Robotics.

⁰In the current implementation of the HSBOF algorithm, many particles are still allocated to irrelevant areas, since no specific representation models are associated to dataless areas. Moreover, if the filtered low level representation can directly be used for various applications (for example mapping process, short-term collision risk assessment [31], [48], etc), the retrospective object level analysis by dynamic grid segmentation can be computationally expensive and subjected to some data association errors.

⁰Nanoelec Technological Research Institute (Institut de Recherche Technologique Nanoelec)

⁰National Research Agency (Agence Nationale de la recherche)

The *CMCDOT* approach introduces a drastic change in the underlying formal expressions: instead of directly filtering the occupancy data, we have added *hidden states* for representing what is currently present in a cell. Then, the occupancy distribution can then be inferred from those hidden states. Besides presenting a clear distinction between static and dynamic parts, the main interest of this new approach is to introduce a specific processing of dataless areas, excluding them from the velocity estimation (and consequently optimizing the processing of the dynamic parts) and disabling their temporal persistence (which is used to generate estimation bias in newly discovered areas). This updated formalism also enables the introduction of an appropriate formal model for the particle initialization and management (which was previously more isolated).

Another important added feature is the automatic segmentation of the dynamic parts of the occupancy grid, according to its shapes and dynamics. While the *CMCDOT* tracks spatial occupancy in the scene without object segmentation, Detection and Tracking of Moving Objects (DATMO) is often required for high level processing. A standard approach would be to analyse the *CMCDOT* outputs, to apply a clustering algorithm on the occupancy grid (enhanced by velocities), and to use those clusters as potential object level targets. This clustering can turn out to be computationally expensive, considering the grid dimensions and the size and complexity of the dynamic particle model. The basic idea of our new approach is to exploit the particle propagation process within the *CMCDOT*: the way particles are resampled can lead to the wanted segmentation after a number of time steps. After initialization, at each step, the particles that correctly fit the motion of a dynamic object are multiplied, those which do not are forgotten. In a few steps, the best particles propagate in the object, and the object motion is fully described by a set of particles deriving from a common particle root. By marking each particle at the initialization step with a unique identification number, all the dynamic areas which are coherent in term of space and motion are marked after few iterations. The convergence of those markers is fastened by additional rules.

Sequence	Grid size (m)	HSBOF	CMCDOT
Highway	20x70	76.9%	23.5%
Semi-Urban	30x60	89.3%	46.7%
City Center	30x60	93.2%	40.1%

Figure 2. Estimation of irrelevant particle allocation ratio.

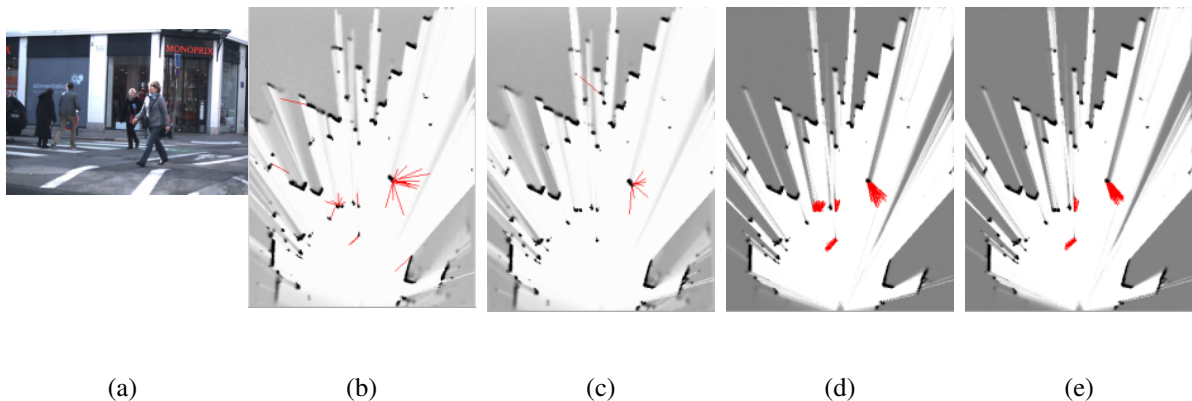


Figure 3. Results of the HSBOF with 262144 and 32768 particles (b) and (c), and of the CMCDOT with the same number of particles (d) and (e). Red segments represent the average estimated per-cell velocity. They show that the CMCDOT is more accurate and still manages to track most of the moving pedestrians (even with a severely reduced number of samples), whereas the HSBOF loses track of almost all objects.

Experimental results showed that the insertion of an "unknown" state in the model leads to a better distribution of dynamic samples on observed areas (see figure 2) and also allows us to be more reactive and accurate on the velocity distributions, while requiring less computing power (see figure 3).

The intrinsic clustering approach has also been tested on real road data, showing promising results in real-time tracking of moving objects, regardless of their type. The method could be improved by managing split-and-merge events that can occur in complex urban environment (see figure 4).

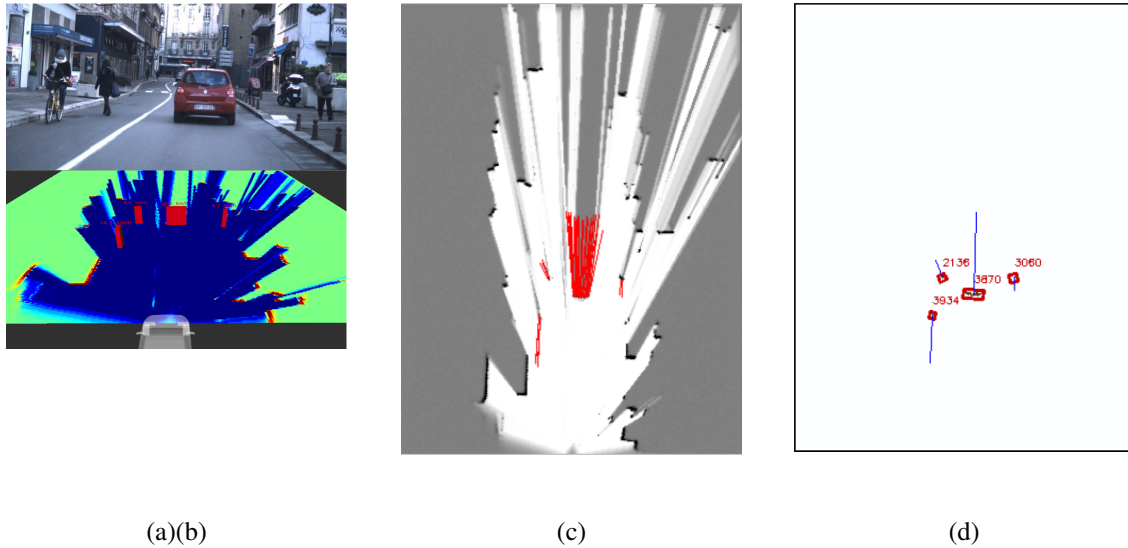


Figure 4. Result of the dynamic objects clustering. (a) Camera image; (b) 3D view of the grid with detected objects; (c) resulting occupancy grid with velocity; (d) extracted dynamic objects (red boxes) with velocity (blue segments) and id.

7.2.2. Multimodal dynamic objects classification

Participants: Amaury Nègre, Jean-Alix David.

The method described in section 7.2.1 allows to obtain a list of dynamic objects and to track each object over time. In order to increase the level of representation of the environment, we have developed a method to classify detected objects using both the camera images and the occupancy grid representation estimated by the *CMCDOT*. For each detected object, the bounding box of the object is projected in the camera image and a local image is extracted from the camera. Jointly, we can extract a patch from the occupancy grid around the dynamic object position. The extracted camera image and the occupancy grid patch can then be used as the input of a Deep Neural Network (DNN) to identify the class of the object. The DNN we designed is a combination on two classic neural networks, the "ImageNet" Convolutional Neural Networks [35] for the camera image input and the "LeNet" [37] for the occupancy grid input (see fig 5).

To train and evaluate the model, a dataset has been created from the data recorded with the Lexus platform. We extracted the camera images and the occupancy grid for each object detected by the *CMCDOT* module, then we manually annotated the object class among "pedestrian", "crowd", "car", "truck", "two-wheelers" and "misc" categories. The resulting dataset contains more than 100000 camera images & occupancy grid pairs. The training process and the classification module has been done by using the open source library caffe [33]. An example of the obtained results is shown on fig 6. The percentage of good classification is greater than 90% on our evaluation dataset.

7.2.3. Visual Map-Based Localisation with OSM

Participants: Jean-Alix David, Amaury Nègre.

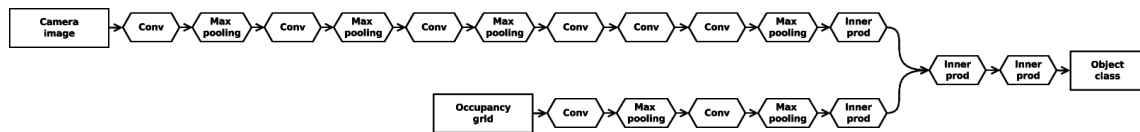


Figure 5. Deep neural network used for obstacle detection.

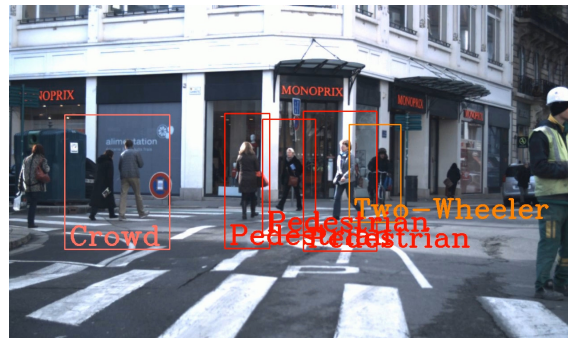


Figure 6. Result of the dynamic obstacle classification.

This module aims to improve both the global localization provided by the GPS⁰ and the lane-relative localization information estimated by a lane tracker by combining their mutual strengths. The idea is to detect lane markings on the road using a camera, and then to compare the extracted lines with those stored in the map. This is done using the ICP⁰ algorithm. This work is described in a confidential Toyota project report entitled *Real Traffic Data Acquisition and Risk Assessment Experiments*.

7.2.3.1. The map

Our solution is based on a post-processed OSM⁰ map shown on figure 7. Typically, these maps contain information on the roads and lanes, but contain no information about lane markers on the ground. Thus, we ran a semi-manual process to complete the existing maps with information about the number and type of markers.

New data are stocked in a local server. Requests can be sent to this server to fetch map data using HTTP protocol.

7.2.3.2. Line detection

The line extraction is done using ridge detection on a top-down view of the camera image. Only one monocular camera is used, as it is an inexpensive sensor, and needs only to be calibrated once. The line detection is based on an algorithm using Laplacian to extract ridges of the monochrome image. The algorithm is implemented for parallelized calculation using CUDA on a GPU, for an improved performance. Figure 8 shows the results of the ridge detector.

⁰Global Positioning System

⁰Iterative Closest Point

⁰OpenStreetMap

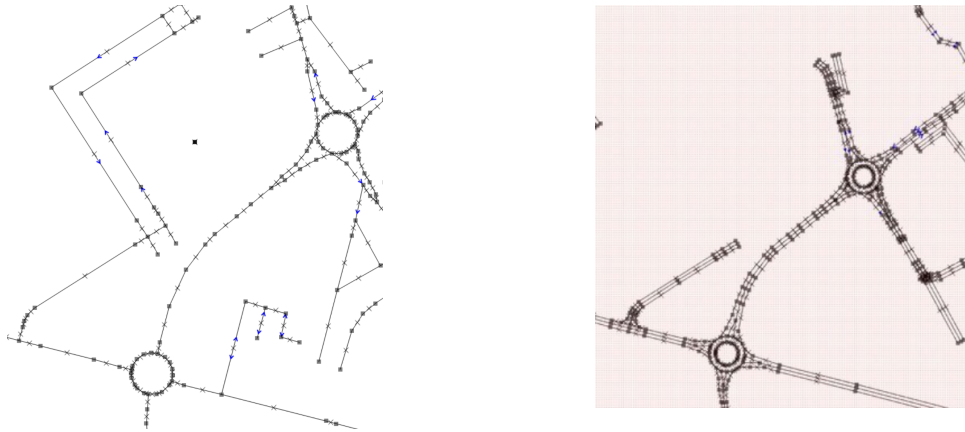


Figure 7. Data conversion. (a) Raw OSM data, a line represents a road and it is not possible to see the lanes. (b) Modified OSM data, with lane markings.

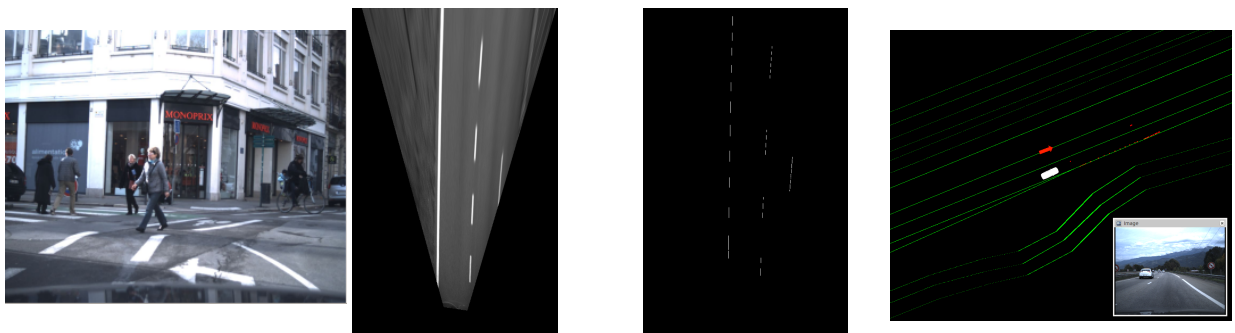


Figure 8. Ridges detection: (a) Input image (b) Projected image (c) Detected ridges. ICP correction on highway (d)

7.2.3.3. ICP-based line matching

The extracted lines are matched and aligned with the map using the ICP algorithm to improve the localization of the vehicle. The ICP algorithm iteratively minimizes the total alignment error between the points detected as ridges and the segments of line extracted from the map. Finally, figure 8 .d shows how the algorithm can correct the vehicle localization. The algorithm is able to accurately track the orientation and position. However, the lateral displacement may be off by a multiple of the lane width, depending on how the algorithm has been initialized. In practice, this effect is often mitigated due to the existence of single-lane roads such as highway entrances.

The results are very promising on highways, but the algorithm has a lower performance on other types of roads, mostly due to irregularities.

7.2.4. Integration of Bayesian Perception System on Embedded Platforms

Participants: Tiana Rakotovoao, Christian Laugier, Diego Puschini(cea Dacle), Julien Mottin(cea Dacle).

Safe autonomous vehicles will emerge when comprehensive perception systems will be successfully integrated into vehicles. However, our Bayesian Perception approach requires high computational loads that are not supported by the embedded architectures currently used in standard automotive ECUs.

To address this issue, we first explored new embedded hardware architecture credible for the integration of OGs⁰ into autonomous vehicles [19]. We studied in particular recent emerging many-core architectures, which offer higher computing performance while drastically reducing the required power consumption (typically less than 1W). In such architectures, the computation of OGs can be divided into several independent tasks, executed simultaneously on separated processing core of a many-core.

Experiments were conducted on data collected from urban traffic scenario, produced by 8 LIDAR layers mounted on the Inria-Toyota experimental Lexus vehicle. These experiments demonstrate that the many-core produces OGs largely in real-time: 6 time faster than the sensor reading rate.

Besides, we also proposed a mathematical improvement of the OG model, for performing multi-sensor fusion more efficiently than the standard approach presented in [29]. In our approach, the fusion of occupancy probabilities requires fewer operations. This model improvement makes it possible the implementation of OG-based multi-sensor fusion on simple hardware architectures. This perspective applies to microcontroller, ASICs or FPGAs which are more and more present in computing platforms recently present on the automotive market.

7.2.5. Experimental Vehicle Renault ZOE

Participants: Nicolas Turro (sed), Jean-François Cuniberto (sed), Procópio Silveira-Stein, Amaury Nègre, Lukas Rummelhard, Jean-Alix David, Christian Laugier.

7.2.5.1. Experimental Vehicle Renault ZOE

In the scope of the *Perfect* projet of the IRT nanoelec, we have started to develop in 2014, an experimental platform based on an equipped Renault Zoe. The development of this platform has been pursued in 2015.

The vehicle has been enhanced with a tablet to display the new HMI⁰, figure 9 (a) illustrates. The HMI displays the detected dynamic objects over the camera image and the graph of collision risk at different time horizon.

New experiments have also been designed to test the perception algorithms and the recent implementation of the collision risk alert. These experiments simulate collisions with people using a fabric mannequin, as shown on figure 9 (b), and an inflatable ball.

⁰Occupancy Grids

⁰Human Machine Interface

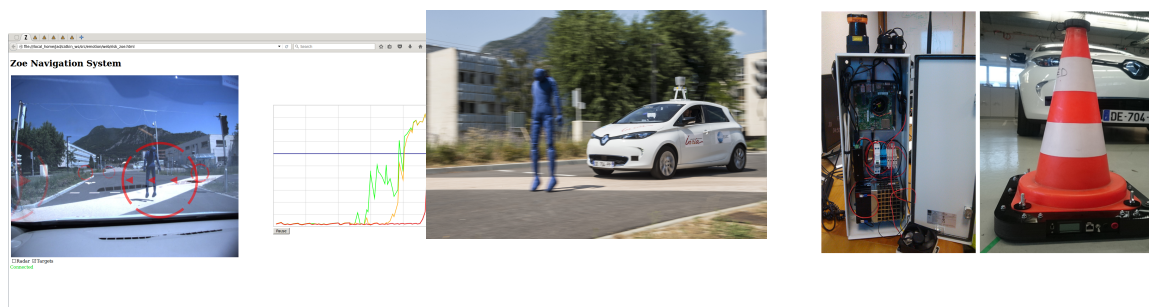


Figure 9. (a) Display of the HMI (b) Collision simulation with a mannequin (c) On left: picture of the smartbox, on right: picture of the cone.

Finally, we have also developed two movable devices in order to enhance V2X⁰ communication experiments (see figure 9 (c)):

1. A movable communicating cone equipped with a GPS and a V2X communication box, which broadcast its position to near V2X listeners.
2. A movable smartbox equipped with a GPS, a V2X communication box, a LIDAR sensor and a Nvidia Tegra K1 board. The *CMCDOT* algorithm is implemented on it, and the detected objects are broadcasted to other communicating devices. The smartbox can be mounted on another vehicle or be placed as part of a static infrastructure. Both are alimented by batteries and aim at minimizing their energy consumption.

7.3. Situation Awareness

Participants: Christian Laugier, Alejandro Dizan Vasquez Govea, Procópio Silveira-Stein, David Sierra-Gonzalez, Mathieu Barbier, Stephanie Lefevre(uc Berkeley).

7.3.1. Framework for Motion Prediction and Collision Risk Assessment

Participants: Christian Laugier, Alejandro Dizan Vasquez Govea, Procópio Silveira-Stein, Stephanie Lefevre(uc Berkeley).

For several years, the challenging scientific problem of Motion Prediction, Risk Assessment and Decision-Making in open and dynamic environments has been one of our main research topics (see activity reports of the former e-Motion Inria team-project). Throughout 2016, we have continued this line of work by developing several new frameworks for Motion Prediction and Collision Risk Assessment in complex dynamic scenes involving multiple moving agents having various behaviors.

A first contribution has been the extensive experimental validation in real conditions –together with the University of Berkeley– of our *Intention-Expectation* approach: a high-level approach to risk assessment which avoids the complexity of trajectory-level reasoning while being able to take multi-vehicle interactions into account [9]. These results have also been integrated into a Mooc course at the graduate and undergraduate levels [25]. They have also been presented in several invited talks [24] [21] [22] [23].

Another contribution relies in the implementation of some the proposed models on two experimental vehicles (Lexus and Zoé experimental platforms). As mentioned in section 7.2.5 , several experiments on short-term collision risk assesement have been successfully conducted with these platforms (c.f. [10], [15]). This work will be continued in 2016, in the scope of our ongoing collaborative projects with Toyota, Renault and IRT nanoelec.

⁰Vehicle-to-Vehicle and Vehicle-to-Infrastructure

7.3.2. Planning-based motion prediction for collision risk estimation in autonomous driving scenarios

Participants: David Sierra-Gonzalez, Alejandro Dizan Vasquez Govea, Christian Laugier.

The objective is to develop a collision risk estimation system capable of reliably finding the risk of collision associated to the different feasible trajectories of the ego-vehicle. This research work is done in the scope of the Inria-Toyota long-term cooperation and of the PhD thesis work of David Sierra- Gonzales.

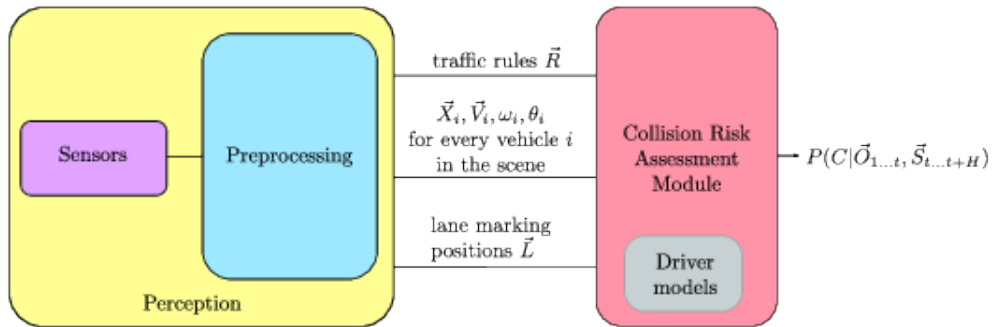


Figure 10. Black box model for the Collision Risk Assessment Module. The number of inputs expected from the perception module is not comprehensive at this point.

Figure 10 shows the black box model of the system. At a given timestep, the system takes the following inputs: the traffic rules in effect; the position, velocity, angular velocity and heading of each vehicle i in the scene; and the position of the lane markings. Thus, at each timestep t we construct an observation vector $1\vec{O}_t = (\vec{R}_t, \{\vec{X}_i, \vec{V}_i, \omega_i, \theta_i\}_t, \vec{L}_t)$ with all the high-level perception inputs, and a state vector \vec{S}_t with only the minimum variables necessary to describe the scene. The proposed system aims to calculate the probability of collision C of the ego-vehicle for a sequence of future states up until a fixed time horizon H . That can be expressed as $P(C | \vec{O}_{1..t}, \vec{S}_{t..t+H})$. This information can then be used by a path-planner to decide upon the safest trajectory.

One key factor for the correct estimation of collision risk is the ability to predict the motion of the dynamic obstacles in the scene, that is, the other drivers. We opt here for a planning-based approach, which assumes that drivers instinctively act as to maximize a reward (or equivalently, minimize a cost). This reward function encodes the preferences of the driver to, for instance, keep a minimum distance with the vehicle in front, drive in the right lane in the highway, or respect the speed limits. Given such a reward function, Markov Decision Processes (MDP) constitute an adequate framework for the motion prediction problem. Moreover, by using Inverse Reinforcement Learning (IRL) algorithms, we can obtain such reward function directly from expert demonstrations (i.e. simply observing how people drive).

At this point, two well-known IRL algorithms ([26], [58]) have been implemented and used to obtain a generic driver model from human demonstrations performed on a highway simulator. This driver model can now be used to predict the future behavior of the dynamic obstacles in the scene.

7.4. Motion-planning in human-populated environment

7.4.1. Planning-based motion prediction for pedestrians in crowded environments

Participant: Alejandro Dizan Vasquez Govea.

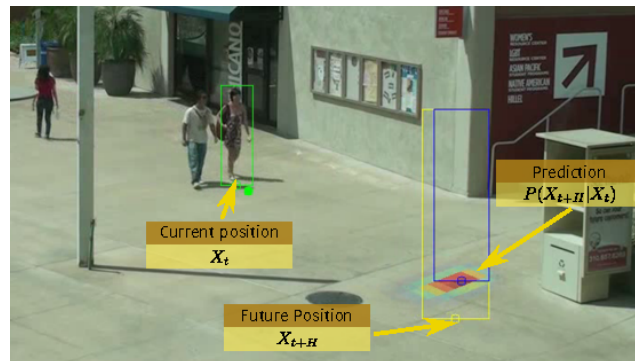


Figure 11. Example of spatiotemporal state prediction for a single agent, indicating its current and future position after 8 sec (green and yellow boxes) as well as the predicted position probability and its maximum value (blue box)

We have also explored the application of motion planning algorithms to the prediction of human motion (Fig. 11). We have proposed a novel planning-based motion prediction approach [12] which addresses the weaknesses of the previous state-of-the-art motion prediction technique [34], namely:

1. *High computational complexity.* This is dealt with by using the Fast Marching Method (FMM) [49] an efficient deterministic planning algorithm which computes the cost-to-go to a given location for every cell of a grid representing the agent's workspace. This grid is then used in a novel goal prediction algorithm and to produce a path-like prediction equivalent to the output of the Markov Decision Processes (MDPs) used by Kitani.
2. *Limited ability to model the temporal evolution along the predicted path:* this is addressed through the use of a velocity-dependent probabilistic motion model which is used to estimate a probability distribution of the future agent's position. This is then fused with a novel cost-based model to produce a full spatiotemporal prediction.
3. *Constant-goal assumption.* We propose a gradient-based goal prediction approach which does not rely on filtering, making it capable of quickly recognizing intended destination changes as they happen.

In our preliminary experiments, the proposed method significantly outperforms the accuracy of Kitani's approach while reducing the computation time by a factor of 30 using a parallel version of our algorithm.

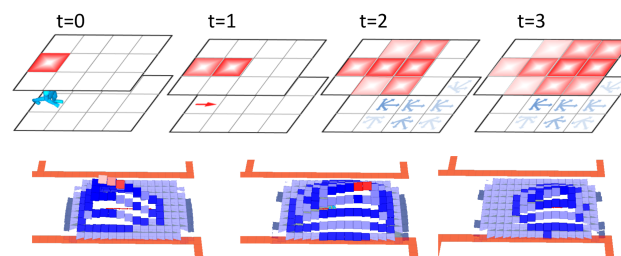


Figure 12. Illustration of (a) the flow-grid model (b) the pheromone model of human motion prediction.

7.4.2. Modeling human-flows from robot(s) perception

Participants: Olivier Simonin, Jacques Saraydaryan, Fabrice Jumel.

To deal with navigation in highly populated environments, eg. flows of humans, we started to investigate the problem of mapping these flows. The challenge is to build such an information from robots perception while they move autonomously to perform their tasks. We also work on predicting humans location from the perceptions and the constructed flow-grid. This led us to define two models : i) a flow-grid mapping computing in each cell the probability to move in each of the k possible directions (illustrated in figure 12 .a), ii) a pheromone-based model allowing to compute the current possible location of humans (flows), see figure 12 .b. We are currently measuring the efficiency of the proposed mapping compared to existing models (which do not model directions). First results will be submitted soon (to IROS 2016).

7.5. Multi-robot Motion-planning in dynamic environments

7.5.1. Benchmarking and extension of multi-robot strategies

7.5.1.1. Exploration of unknown and populated environments

Participants: Olivier Simonin, Nassim Kaldé (phd. Student, Larsen Inria Nancy), François Chapillet (inria Larsen, Nancy), Jan Faigl (ctu, Czech University Of Prague).

Exploration of unknown environment with a group of mobile robots consists mainly to compute a strategy that allows to visit efficiently the area while considering different constraints. These constraints can be trajectory coordination (between robots), presence of humans and limits on time, communication, and computational resources allowed to robots. The exploration problem is related with mapping, surveillance (eg. patrolling) problems. In this context, O. Simonin and P. Lucidarme (University of Angers) published a general article on multi-robot mapping in the magazine Techniques de l'Ingénieur [5] (2015).

Study of frontier-based strategies

In this context, frontier-based approaches looks for an efficient allocation of the navigational goals which must be situated between the known and unknown areas (the frontiers). Goal candidate locations are repeatedly determined during the exploration. Then, the assignment of the candidates to the robots is solved as the task-allocation problem. A more frequent decision-making may improve performance of the exploration, but in a practical deployment of the exploration strategies, the frequency depends on the computational complexity of the task-allocation algorithm and available computational resources. Therefore, we proposed an evaluation framework to study exploration strategies independently on the available computational resources and we reported a comparison of the selected task-allocation algorithms deployed in multi-robot exploration [30]. This work is supported by the French-Czech PHC "Murotex".

Exploration in populated environments

In the context of the Phd of Nassim Kaldé, co-supervised by F. Chapillet (Inria Nancy, Larsen) and O. Simonin (Chroma), we study exploration in populated environments, in which pedestrian flows can severely impact performances. However, humans have adaptive skills for taking advantage of these flows while moving. Therefore, in order to exploit these human abilities, we propose a novel exploration strategy that explicitly allows for human-robot interactions. Our model for exploration in populated environments combines the classical frontier-based strategy with our interactive approach. For this purpose, we proposed an interaction model where robots can locally choose a human guide to follow and define a parametric heuristic to balance interaction and frontier assignments. This model is introduced in publication [3], where we evaluate to which extent human presence impacts the exploration model in terms of coverage ratio, travelled distance and elapsed time to completion. A simulator, based on V-REP and illustrated in figure 13 .a, has been developed to conduce the experimental measures.

7.5.1.2. Patrolling static and dynamic environments

Participants: Olivier Simonin, Jacques Saraydaryan, Fabrice Jumel, Mihai Popescu, Herve Rivano (inria Urbanet).

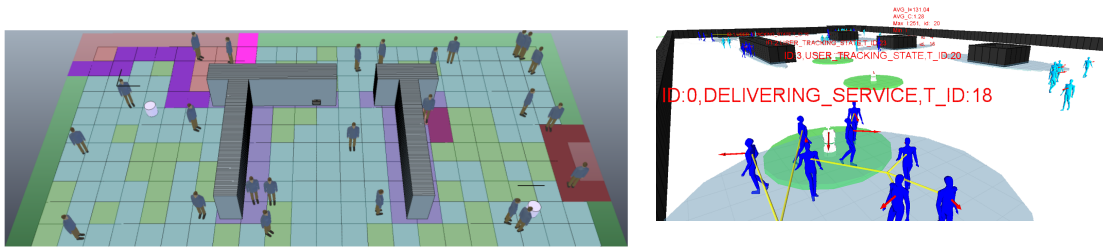


Figure 13. (a) Simulator to study exploration in populated environment based on V-REP (b) Simulator for dynamic patrolling of people based on PedSim.

Patrolling moving people

In the context of service robotics, we address the problem of serving people by a set of collaborating robots, that is to deliver regularly services to moving people. We re-defined this problem as a dynamic patrolling task, that we called the robot-waiters problem, where robots have to regularly visit all the moving persons. In the publication [11], we proposed different criteria and metrics suitable to this problem, by considering not only the time to patrol all the people but also the equity of the delivery. We proposed and compared four algorithms, two are based on standard solutions to the static patrolling problem and two are defined according the specificity of patrolling moving entities. In order to limit robot traveled distances, the last approach introduces a clustering heuristic to identify groups among people. To compare algorithms and to prepare real experiments we developed a simulator combining a pedestrian model (PedSim) and a robotic model, illustrated in figure 13 .b. Experimental results show the efficiency of the specific new approaches over standard approaches. We also analysed the influence of the number of robots on the performances, for each approach.

We are currently developing new algorithms using the mapping and prediction of human flows based on the work presented in section 7.4.2 .

Patrolling WSN

In the multi-robot patrolling context, we investigated the problem of visiting regularly a set of fixed sensors by computing single-cycles on the graph formed by the WSN (Wireless sensors network). We set this problem as a graph covering with bounded hamiltonian cycles (in the M2R internship of Mihai-Ioan Popescu, now continuing as PhD student in Chroma). After giving insights of NP-hardness, we proposed a generic heuristic algorithm for solving the GCBHC. It works in two steps: the first one partitions the vertices, the second one computes hamiltonian cycles on each partition. We adapted the classic Multilevel Subgraph Partitioning algorithm to the specific requirements yielded by the networking metrics. To avoid the high complexity of this algorithm, we proposed another heuristic which exploits the geometric structure of the graph, the North-Eastern Neighbour heuristic. We implemented two classic hamiltonian cycle heuristics, one is based on Minimum Spanning Trees computations and the other on Christofides algorithm. Comparisons on randomly-generated graphs showed that the Christofides algorithm computes shorter cycles. An article presenting this work has been written and will be submitted soon.

7.5.2. Anytime algorithms for multi-robot cooperation

7.5.2.1. Observation of complex scenes

Participants: Olivier Simonin, Jilles Dibangoye, Laetitia Matignon (Iris), Christian Wolf (Iris), Jonathan Cohen (internship), Stefan Chitic.

Solving complex tasks with a fleet of robots requires to develop generic strategies that can decide in real time (or time-bounded) efficient and cooperative actions. This is particularly challenging in complex real environments. To this end, we explore anytime algorithms and adaptive/learning techniques.

The INSA BQR project "Crome"⁰, led by O. Simonin, motivated the exploration of the joint-observation of complex (dynamic) scenes by a fleet of mobile robots. In our current work, the considered scenes are defined as a sequence of activities, performed by a person in a same place. Then, mobile robots have to cooperate to find a spatial configuration around the scene that maximizes the joint observation of the human pose skeleton. It is assumed that the robots can communicate but have no map of the environment and no external localisation.

To attack the problem, in cooperation with colleagues from vision (C. Wolf, Liris), we proposed an original concentric navigation model allowing to keep easily each robot camera towards the scene (see fig. 14 .a). This model is combined with an incremental mapping of the environment in order to limit the complexity of the exploration state space. We have also defined the marginal contribution of each robot observation, to facilitate stability in the search, while the exploration is guided by a meta-heuristics. We developed a simulator (fig. 14 .b) that allows to compare the variants of the approach and to show its features such as adaptation to the dynamic of the scene and robustness to the noise in the observations. Preliminary results have been presented in [8].

We have also developed an experimental framework, using Turtlebot2 robots, presented in figure 14 .c. Experimental measures and validation are in progress.

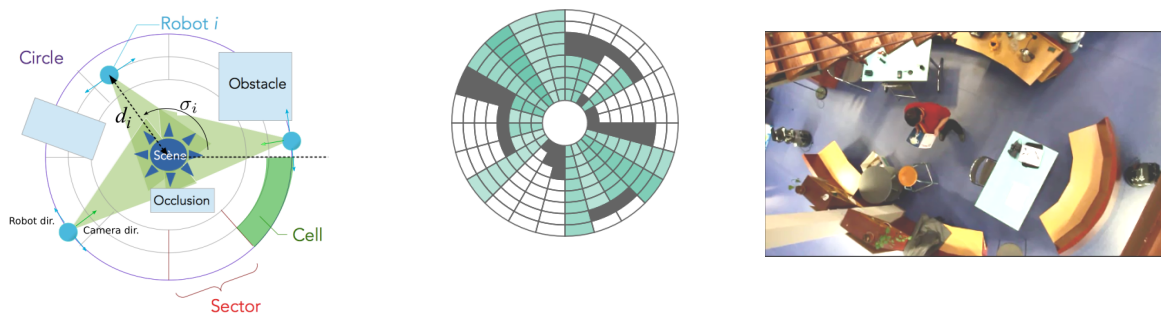


Figure 14. Illustrations (a) Concentric navigation model, (b) Simulator and (c) experimental setup with Turtlebot 2.

7.5.2.2. Middleware for multi-robot systems deployment

Participants: Stefan Chitic, Julien Ponge (citi, Dynamid), Olivier Simonin.

Multi-robots systems (MRS) require dedicated tools and models to face the complexity of their design and deployment (there is no or very limited tools/middleware for MRS). In this context, we addressed the problem of neighbors and service discovery in an ad-hoc network formed by a fleet of robots. Robots needs a protocol that is able to constantly discover new robots in their coverage area. This led us to propose a robotic middleware, SDfR, that is able to provide service discovery. This protocol is an extension of the Simple Service Discovery Protocol (SSDP) used in Universal Plug and Play (UPnP) to dynamic networks generated by the mobility of the robots. Even if SDfR is platform independent, we proposed a ROS integration in order to facilitate the usage. We evaluated a series of overhead benchmarking across static and dynamic scenarios. Eventually, we experimented some use-cases where our proposal was successfully tested with Turtlebot 2 robots. Results have been presented to the national conference CAR'2015 and will appear in the international conference ICAART 2016 (accepted).

⁰Coordination d'une flottille de robots mobiles pour l'analyse multi-vue de scènes complexes

7.5.3. Sequential decision-making under uncertainty

Sequential decision-making under uncertainty is a core area of artificial intelligence, optimization, operations research, machine learning, and robotics. It involves one or multiple decision makers (or agents or robots) reasoning about the course of actions to achieve collective or self-interested goals while accounting both for the outcomes of current decisions and for future decision-making opportunities. Markov models (e.g., Markov decision processes and Markov games) have emerged as normative frameworks for optimizing decision under uncertainty. These models encompass a wide range of real-world applications: controlling intelligent vehicles; optimizing the production and distribution of energy resources; protecting endangered species; making telecommunication protocols faster and safer; monitoring and assisting elderly patients at home; designing robotic exploration technologies for search and rescue; but also many other applications. Decentralized partially observable Markov decision processes have emerged as the fundamental model to address multiple decision makers' decision-theoretic planning and learning problems. In that direction, we investigate generic, highly scalable and adaptable planning and learning algorithms to apply eventually in multi-robot planning tasks.

7.5.3.1. Structural results for cooperative decentralized control problems

Participants: Jilles S. Dibangoye, Olivier Simonin, Olivier Buffet (inria Nancy, Ex Maia Team), Mamoun Idrissi (internship, Insa de Lyon).

The intractability in cooperative, decentralized control problems is mainly due to prohibitive memory requirements in both optimal policies and value functions. The complexity analysis has emerged as the standard method to estimating the memory needed for solving a given computational problem, but complexity results may be somewhat limited. Our work [13] introduces a general methodology, called the structural analysis, for the design of optimality-preserving concise policies and value functions, which will eventually lead to the development of efficient theory and algorithms. For the first time, we showed that memory requirements for policies and value functions may be asymmetric, resulting in cooperative, decentralized control problems with exponential reductions in memory requirements. To apply this theoretical in robotics, we investigate during M. Idrissi's internship the robotic coverage of unknown areas.

7.5.3.2. State-of-the-art algorithms for optimally solving Dec-POMDPs

Participants: Jilles S. Dibangoye, Christopher Amato (univ. New Hampshire), Olivier Buffet (inria Nancy, Ex Maia Team), François Charpillat (inria Nancy, Larsen Team), Martin Pugno (master Student, U. Claude Bernard Lyon).

Decentralized partially observable Markov decision processes (Dec-POMDPs) provide a general model for decision-making under uncertainty in cooperative decentralized settings but are difficult to solve optimally (NEXP-Complete). As a new way of solving these problems, we introduced the idea of transforming a Dec-POMDP into a continuous-state deterministic MDP with a piecewise-linear and convex value function. This approach makes use of the fact that planning can be accomplished in a centralized offline manner while execution can still be decentralized. This new Dec-POMDP formulation, which we call an occupancy MDP, allows powerful POMDP and continuous-state MDP methods to be used for the first time. To provide scalability, we refine this approach by combining heuristic search and compact representations that exploit the structure present in multi-agent domains, without losing the ability to converge to an optimal solution. In particular, in [14], we introduce a feature-based heuristic search value iteration (FB-HSVI) algorithm that relies on feature-based compact representations, point-based updates, and efficient action selection. However, scalability remains limited when the number of agents or problem variables becomes large. To overcome this limitation, we show that, under certain separability conditions of the optimal value function, the scalability of this approach can increase considerably. This separability is present when there is the locality of interaction between agents, which can be exploited to improve performance. A theoretical analysis demonstrates that FB-HSVI terminates in finite time with an optimal solution. We include an extensive empirical analysis using well-known benchmarks, thereby confirming that our approach provides significant scalability improvements compared to the state of the art. We push even further the envelope, during Martin's internship, assuming we only have access to an incomplete model of the world. This more realistic assumption that would ease application to robotics leads us directly to learning algorithms inspired from FB-HSVI.

7.5.3.3. *Distributed projected gradient-descent algorithm applied to smart grids*

Participants: Jilles S. Dibangoye, Arnaud Doniec (uria – Ecole Des Mines de Douai, France), H. Fakham, F. Colas And X. Guillaud (I2ep – Arts Et Métiers Paristech, France).

In a smart grid context, the increasing penetration of embedded generation units leads to a greater complexity in the management of production units. In this work, we focus on the impact of the introduction of decentralized generation for the unit commitment (UC) problem. Unit commitment problems consist in finding the optimal schedules and amounts of power to be generated by a set of generating units in response to an electricity demand forecast. While this problem has received a significant amount of attention, classical approaches assume that these problems are centralized and deterministic. However, these two assumptions are not realistic in a smart grid context. Indeed, finding the optimal schedules and amounts of power to be generated by multiple distributed generator units is not trivial since it requires to deal with distributed computation, privacy, stochastic planning, etc. Our contribution focuses on smart grid scenarios where the main source of complexity comes from the proliferation of distributed generating units. In solving this issue, we consider distributed stochastic unit commitment problems. In [2], we introduce a novel distributed gradient descent algorithm which allows us to circumvent classical assumptions. This algorithm is evaluated through a set of experiments on real-time power grid simulator.

DAHU Project-Team

6. New Results

6.1. Specification and verification of data-driven systems

Process-centric views of data-driven workflows. Declarative, data-aware workflow models are becoming increasingly pervasive. While these have numerous benefits, views describing valid sequences of tasks are also useful to provide stake-holders with high-level descriptions of the workflow. In [23], we study the problem of recovering process-centric views from declarative, data-aware workflow specifications. The views are most naturally specified by finite-state transition systems describing regular languages. The results characterize when process-centric views of artifact systems are regular, with both linear and branching-time semantics.

Complexity in counter systems and in proof systems. The static analysis of queries on XML trees and data streams relies in a majority of cases on decision procedures expressed in terms of formal systems like counter systems or proof systems. For instance, two-variables first-order data queries on words can be related to reachability in vector addition systems (VAS), and the same queries on trees to reachability in a branching extension of VAS. We have fundamental results on the computational complexity of these problems, including the first explicit upper bound for reachability in VAS [24] and the best known lower bound for reachability in branching VAS [17] (where it is currently unknown whether reachability is decidable at all). We have furthermore defined a first sequent calculus for a modal data logic [29] as preliminary groundwork for the ANR PRODAQ project on proof systems for data queries.

6.2. Query processing for the Web

Query languages for graph databases. Graph-structured data on the Web can be found in emerging applications such as *RDF* and *linked data*, or *social networks*. Classical database languages are not suitable to query such data, essentially because they do not allow to (easily) express simple connectivity queries, which are the basic building block in graph navigation. We use Regular Path Queries, computing pairs of nodes reachable via a path satisfying a regular expression. We have tackled the problem of answering queries over graph databases which are available only through a given set of views. We have shown that in the “asymptotic case”, i.e. when the query is large enough relative to the view definition, it is decidable whether the view determines the query [22].

6.3. Distributed knowledge base.

Webdamlog The Webdamlog language is an extension of datalog to the distributed context, with *delegation* as the main novelty. A summary of the project was presented in [20].

We introduced an access control mechanism based on provenance in [26]. This access control is designed for a distributed setting. Peers getting data are also willing to enforce access control on that data, so that the owner of the data keeps some control over it when the data is passed around in the network. A second version of Webdamlog was developed in 2015 at Drexel, primarily by Vera Moffit also as part of her thesis (in collaboration with S. Abiteboul). It includes access control mechanism.

DEFROST Team

7. New Results

7.1. Aggregate Constraints for Virtual Manipulation with Soft Fingers

In this work, we propose a new formulation of contact and friction laws, in the context of virtual grasping. The work allows to reduce the number of contact and friction constraints, using volume interpenetration measure, instead of interpenetration distance. The work has been conducted in collaboration with Antony Talvas and Maud Machal (Inria Hybrid Team, Rennes) and Miguel Otaduy (URJC Madrid). It has been presented at the conference IEEE VR and published in the journal TVCG [5].

7.2. Haptic Rendering of Hyperelastic Models with Friction

We have reached an important milestone with this work: we have merge two important research tracks of these last years: On one hand, haptic rendering of friction contact between deformable objects ; on the other hand, real-time simulation of hyperelastic objects (particularly to simulate soft-tissues). This work has been conducted in collaboration with Hadrien Courtecuisse (Inria team Mimesis) and Hervé Delingette (Inria team Asclepios) [6]

7.3. Augmentation of Elastic Surfaces with Self-Occlusion Handling

In this work, we propose to recover the 3D shape and to augment elastic objects with self-occlusions handling, using only single view images. Shape recovery from a monocular video sequence is an underconstrained problem and many approaches have been proposed to enforce constraints and resolve the ambiguities. State-of-the-art solutions enforce smoothness or geometric constraints, consider specific deformation properties such as inextensibility or resort to shading constraints. We propose a real-time method that uses a mechanical model and that is able to handle highly elastic objects. The problem is formulated as an energy minimization problem accounting for a non-linear elastic model constrained by external image points acquired from a monocular camera. This method prevents us from formulating restrictive assumptions and specific constraint terms in the minimization. In addition, we propose to handle self-occluded regions thanks to the ability of mechanical models to provide appropriate predictions of the shape. This result has been published in the journal TVCG [2] and has been extended to handle cuttable objects and has been published as a SIGGRAPH poster [12].

7.4. Real-time control of soft-robots using asynchronous finite element modeling

Finite Element analysis can provide accurate deformable models for soft-robots. However, using such models is very difficult in a real-time system of control. In this study, we introduce a generic solution that enables a high-rate control and that is compatible with strong real-time constraints. From a Finite Element analysis, computed at low rate, an inverse model of the robot outputs the setpoint values for the actuator in order to obtain a desired trajectory. This inverse problem uses a QP (quadratic-programming) algorithm based on the equations set by the Finite Element Method. To improve the update rate performances, we propose an asynchronous simulation framework that provides a better trade-off between the deformation accuracy and the computational burden. Complex computations such as accurate FEM deformations are done at low frequency while the control is performed at high frequency with strong real-time constraints. The two simulation loops (high frequency and low frequency loops) are mechanically coupled in order to guarantee mechanical accuracy of the system over time. Finally, the validity of the multi-rate simulation is discussed based on measurements of the evolution in the QP matrix and an experimental validation is conducted to validate the correctness of the high-rate inverse model on a real robot. [8]

7.5. Domain decomposition approach for FEM quasistatic modeling and control of Continuum Robots with rigid vertebrae

This study focuses on a new method dedicated to the modeling and control of Continuum Robots, based on the Finite Element Method (FEM) using quasi-static assumption. The modeling relies on a discretization of the continuum robots using 6 DoFs Frames along the structure of the robot that is compatible with the modeling of a sequence of rigid vertebrae. When the robot's structure relies on rods with constant sections, internal forces are computed with beam elements, placed between two adjacent frames, that applies forces and torques. In the opposite, when the robot is composed of a complex shape deformable backbone separated by the rigid vertebrae, a domain decomposition strategy is used to obtain an equivalent stiffness between two vertebrae using volumetric FEM. In both cases, for solving the whole robot model and inverting it in real-time, the numerical method takes advantage of the serial nature of continuum robots, using a Block-Tri-Diagonal solver. The factor of improvement in the computation time reaches several order of magnitude compared to a classical FEM model, while keeping a good precision. The method has also been implemented and tested on a real pneumatic CBHA trunk designed by Festo Robotics and some complementarity examples have been generated numerically.[10]

DREAM Project-Team

7. New Results

7.1. Simulator-based decision support

Participants: Philippe Besnard, Marie-Odile Cordier, Anne-Isabelle Graux, Christine Largouët, Véronique Masson, Laurence Rozé.

7.1.1. Ecosystem model-checking for decision-aid

Former studies of ecosystem modelling have concentrated on temporal modelling. In recent studies we have focussed on the formalization of spatial diffusion of a prey-predator trophic network composed of weeds and ground beetle. For this purpose, an approach coupling landscape representation and population models has been used. A reaction-diffusion model was developed through the synchronization ability of timed-automata. The agronomical rules of beetle migration and weeds diffusion have been translated into communications between timed automata. Landscapes have been simulated and can be evaluated thanks to landscape-metrics distance. The optimization aims to maximize the ground beetle abundance while minimizing the use of pesticides. The model obtained in this first study is quite complex but preliminary results are being studied.

7.1.2. Controller synthesis for optimal strategy search

Similarly to previous work, this approach relies on a qualitative model of a dynamical system. The problem consists in finding a strategy in order to help the user achieving a specific goal. The model is now considered as a timed game automata expressing controllable and uncontrollable actions. The strategy represents the sequence of actions that can be performed by a user to reach a particular state (in case of a reachability problem for instance). A first approach based on a "generate and test" method has been developed for the marine ecosystem example [86].

Recently, we generalized the work of Yulong Zhao applied in the context of a dairy production system [87] to the planning domain. The planning task consists in selecting and organizing actions in order to reach a goal state in a limited time and in an optimal manner, assuming actions have a cost. We propose to reformulate the planning problem in terms of model-checking and controller synthesis on interacting agents such that the state to reach is expressed using temporal logic. We have chosen to represent each agent using the formalism of Priced Timed Game Automata (PTGA). PTGA is an extension of Timed Automata that allows the representation of cost on actions and uncontrollable actions. Relying on this domain description, we define a planning algorithm that computes the best strategy to achieve the goal. This algorithm is based on recognized model-checking and synthesis tools from the UPPAAL suite. The expressivity of this approach is evaluated on the classical *Transport Domain* which is extended in order to include timing constraints, cost values and uncontrollable actions. This work has been implemented and performances evaluated on benchmarks.

7.1.3. A datawarehouse for simulation data

In previous work we have proposed a datawarehouse architecture to store the huge data produced by deep agricultural simulation models [50]. This year, we have worked on hierarchical skyline queries to introduce skyline queries in a datawarehouse framework. Conventional skyline queries retrieve the skyline points in a context of dimensions with a single hierarchical level. However, in some applications with multidimensional and hierarchical data structure (e.g. data warehouses), skyline points may be associated with dimensions having multiple hierarchical levels. Thus, we have proposed an efficient approach reproducing the effect of the OLAP operators "drill-down" and "roll-up" on the computation of skyline queries [52]. It provides the user with navigation operators along the dimensions hierarchies (i.e. specialize / generalize) while ensuring an online calculation of the associated skyline.

Anne-Isabelle Graux, on leave from INRA (National Institute for Agronomical Research), is working on an adaptation and extension of this method for storing the simulation results of a comprehensive farm model named MELODIE [53]. The new datawarehouse will enable the analysis of simulation results within dynamic preferences, related to grassland management for instance, for identifying the data satisfying the best compromises with respect to possibly inconsistent criteria.

7.1.4. Post-mining classification rules

We consider sets of classification rules with quantitative and qualitative attributes inferred by supervised machine learning, as in the framework of the Sacadeau project. Our aim is to improve the human understanding of such sets of rules. First, we consider quantitative attributes in rules that often contain too many intervals which are difficult to interpret. We propose two algorithms to merge some of these intervals in order to get more understandable rules. These algorithms take into account the final rule quality. We are also working on formalizing what could be the quality of a set of rules. There are lots of studies about the quality of one rule but very few about the quality of the whole set of rules and this is still an issue.

7.2. Data Mining

Participants: Marie-Odile Cordier, Yann Dauxais, Serge Vladimir Emteu Tchagou, Clément Gautrais, Thomas Guyet, Yves Moinard, Benjamin Negrevergne, René Quiniou, Laurence Rozé, Alexandre Termier.

7.2.1. Sequential pattern mining with intervals

In previous work, we developed a framework for sequential pattern mining with intervals [3]. It has been applied in various application (care-pathways, customer relationship management databases [35], etc.).

This year we explored chronicle mining algorithms for mining care-pathways (see section 9.1.1, for an applicative context). Chronicles are alternative patterns for representing temporal behaviors [58]. A chronicle can be briefly defined as a set of events linked by constraints indicating the minimum and maximum time elapsed between two events. A care-pathway contains point-based events (e.g. surgery) and interval-based events (e.g. drug exposures). A chronicle can express such a complex temporal behaviour, for instance: *The patient was exposed to a drug X between 1 and 2 years, he met his doctor between 400 to 600 days after the beginning of the exposure and, finally, he was hospitalized.*

The first algorithm we worked on [23] is an adaptation of existing chronicle mining algorithms [55], [63] to mine the complete set of frequent chronicles from a collection of care-pathways. This algorithm uses the search-space browsing strategy of HDCA [55] and the support evaluation of CCP-Miner [63]. As the complete set of chronicle is huge, we also proposed an incomplete algorithms based on the original simplifications of [58]. These algorithms were implemented and evaluated on real and simulated datasets.

We also investigated discriminant chronicles mining which consists in extracting the chronicles that are α times more frequent in a database \mathcal{D}_+ than in a database \mathcal{D}_- . Mining discriminant chronicles is very useful to discover the features of care-pathways that are related, for instance, to a specific disease. Our approach has been implemented and is under evaluation.

7.2.2. Multiscale segmentation of satellite image time series

Satellite images enable the acquisition of large-scale ground vegetation information. Images have been recorded for several years with a high acquisition frequency (one image every two weeks). Such data are called satellite image time series (SITS). Several articles were published this year and they correspond to past work on algorithms and method to analyse SITS.

In [11], we presented a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. The main issue of this approach was the required computation resources (time and memory).

We also explored the supervised classification of SITS using classification trees for time-series [27] by implementing a parallelized version of this algorithm. Next, we explored the adaptation of the object-oriented segmentation to SITS. The object-oriented segmentation is able to segment images based on segment uniformity. We proposed a measure for time-series uniformity and applied the adapted algorithm on large multivariate SITS of Senegal [10].

Third, we presented an supervised approach to extract features from classified satellite images to analyse urban sprawl [28]. In this work, we have satellite images at only two dates, and the objective is to identify characteristics that can foster or prevent changes.

Our satellite images analysis approaches are used in two applicative contexts: understanding urban sprawl and analyzing drought in Senegal. Analysis of urban sprawl was a collaborative work with colleagues in remote sensing, in landscapes analysis and in economical modelling. Our collective contribution was published in a book of the PDD2⁰ program [38]. Analysis of drought in Senegal is a long term collaboration with H. Nicolas (INRA/SAS) that we would like to continue in a collaboration with A. Fall (Université of Dakar) to confront our results with ground observations.

7.2.3. Analysis and simulation of landscape based on spatial patterns

Researchers in agro-environment need a great variety of landscapes to test their scientific hypotheses using agro-ecological models. Real landscapes are difficult to acquire and do not enable the agronomists to test all their hypotheses. Working with simulated landscapes is then an alternative to get a sufficient variety of experimental data. Our objective is to develop an original scheme to generate landscapes that reproduce realistic interface properties between parcels. This approach consists of the extraction of spatial patterns from a real geographic area and the use of these patterns to generate new "realistic" landscapes. It is based on a spatial representation of landscapes by a graph expressing the spatial relationships between the agricultural parcels (as well as the roads, the rivers, the buildings, etc.), in a specific geographic area.

In past years, we worked on the exploration of graph mining techniques, such as gSPAN [85], to discover the relevant spatial patterns present in a spatial-graph. We assume that the set of the frequent graph patterns are the characterisation of the landscape. Our remaining challenge was to simulate new realistic landscapes that reproduce the same patterns.

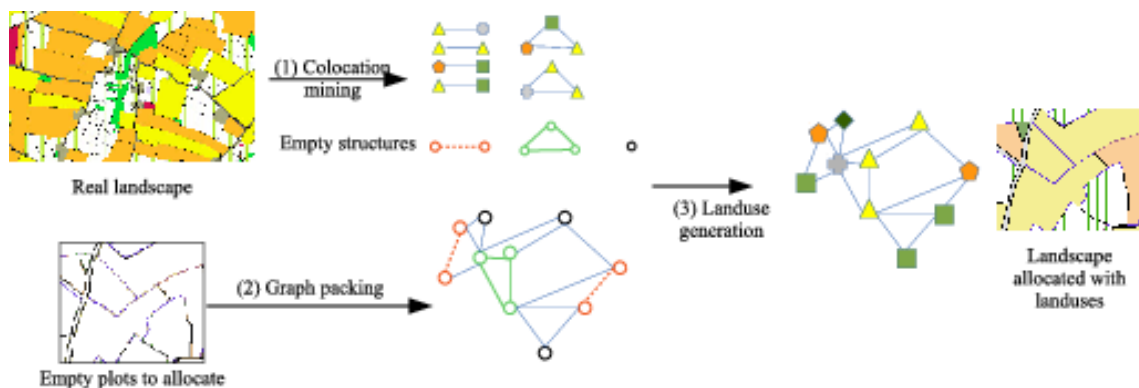


Figure 1. Simulation process in three steps: 1) characteristic graph-patterns mining, 2) graph packing of the cadastral landscape and 3) crop allocation.

⁰PDD2: Paysage Développement Durable/Landscape Sustainable Development

We have formalized the simulation process as a graph packing problem [66]. The process is illustrated by Figure 1. Solving instances of the general graph packing problem has a high combinatorics and no efficient algorithm can solve it. We proposed an ASP program to tackle the combinatorics of the graph packing and to assign the land use considering some expert knowledge. Our approach combines the efficiency of ASP to solve the packing issue and the simplicity of the declarative programming to take into account expert constraints on the land use. Constraints about the minimum surface of crops or about the impossibility of some crops colocation can be easily defined. This work have been presented at the conference RFIA and an extended version has been published in the Revue d'Intelligence Artificielle (RIA) [13].

In addition, we are collaborating with J. Nicolas (EPI Dyliss) to improve the efficiency of our first programs. The improvements are based on symmetry breaking of ASP programs. To this end, we proposed a simplified encoding of the graph patterns using spanning trees and used automorphism detection in graph patterns to automatically encodes symmetry breakings. Intensive evaluation of our encoding shown that this improvement enable to tackle significantly larger graphs than early programs did. This work will be soon submitted to a high ranking conference.

7.2.4. Mining with ASP

In pattern mining, a pattern is considered interesting if it occurs frequently in the data, i.e. the number of its occurrences is greater than a fixed given threshold. As non informed mining methods tend to generate massive results, there is more and more interest in pattern mining algorithms able to mine data considering some expert knowledge. Though a generic pattern mining tool that could be tailored to the specific task of a data-scientist is still a holy grail for pattern mining software designers, some recent attempts have proposed generic pattern mining tools [61] for itemset mining tasks. In collaboration with Torsten Schaub, we explore the ability of a declarative language, such as Answer Set Programming (ASP), to solve pattern mining tasks efficiently. In 2011, Jarvisälo proposed a first attempt devoted to itemset mining [64]. In Dream, we are working on sequential pattern mining, which is known to be more challenging than itemset mining and which has been also recently considered by constraint programming approaches [76].

We have worked on encoding in ASP most of sequential pattern mining tasks: sequences with constraints (gaps, maximum length, etc.), closed/maximal patterns, emergent sequences. Our first result is to show that ASP is suitable for encoding such complex pattern mining tasks. The experimental results show that our purely declarative approach is less efficient than constraint programming approaches [36]. Nonetheless, it is suitable to be blend with intensive knowledge. The challenge is now to show that our ASP framework can extract the meaningful patterns that other approaches loose in the overwhelming amount of sequential patterns.

A first attempt has been done in this direction in collaboration with J. Romero from the University of Potsdam. We used the system ASPRIN to define preferences on patterns. Defining preferences on patterns is also a classical approach to select the most interesting patterns. Some classical preferences on sequential patterns have been defined and the ASPRIN system is used to extract the preferred patterns according to one preference or a combination of preferences (skypatterns [81])

This work will be soon submitted to a high ranking international conference.

7.2.5. Mining time series

Monitoring cattle. Following the lines of a previous work [79], we are working on a method for detecting Bovine Respiratory Diseases (BRD) from behavioral (walking, lying, feeding and drinking activity) and physiological (rumen temperature) data recorded on feedlot cattle being fattened up in big farms in Alberta (Canada). This year, we have especially worked on multivariate sensor data analysis, especially on the evaluation of different combinations of sensors for determining the best configuration and parameter setting. This work was part of Afra Verena Mang's master thesis defended in september 2015 [73]. Two papers are in preparation.

SIFT-based time-series symbolisation Time series classification is an application of particular interest with the increase of such data. Computing the distance between time-series is time consuming. An abstract representation of time-series that accurately approximates distances between time-series and makes easier

their comparison is highly expected. In [17], we proposed a time series classification scheme grounded on the SIFT framework [70] adapted to time series. The SIFTs feed a Bag-of-Words representation of time-series. We have shown that this framework efficiently and accurately classifies time series, despite the fact that BoW representation ignores temporal order.

Mining sequential patterns from multimedia data Analyzing multimedia data to extract knowledge is a challenging problem due to the quantity and complexity of such data. Finding recurrent patterns is one method to structure and segment the data. In a collaboration with the EPI LinkMedia, we have proposed audio data symbolization and sequential pattern mining methods to extract patterns from audio streams. Experiments show this the task is hard and that the symbolization is a critical step for extracting relevant audio patterns [29].

7.2.6. Mining customer data for predicting and explaining attrition

Predicting customer defection in a retail context is difficult because, in most situations, the customer does not leave the store totally (there is no contract break as with banks or phone operators). We have proposed a new pattern model for representing the evolution of an individual customer purchase behavior that enables to early detect and to explain customer attrition. In particular, this model enables the analyst to determine which important kinds of product receives less and less attention from the customer. Thus, this model provides actionable knowledge at an individual scale that lets the retailer trigger targeted marketing actions to counter attrition. A poster has been submitted to the EBDT conference. This work has been performed during Clément Gautrais's master [59] and will be further investigated and extended during his PhD.

7.2.7. Mining energy consumption data

Machine tools in companies consume a lot of energy (before, during and after producing worked pieces). This year, we are beginning to work, with the start-up Energiency, on mining machine tool energy consumption data in order to propose energy savings to the companies. Firstly, we try to determine, according to the analyzed company, which data-mining algorithm should be used and which is the best configuration and parameter setting. Then, we aim to extract actions rules from patterns to help companies to consume less energy.

7.2.8. Trace reduction

One problem of execution trace of applications on embedded systems is that they can grow very large, typically several Gigabytes for 5 minutes of audio/video playback. Some endurance tests require continuous playback for 96 hours, which would lead to hundreds of Gigabytes of traces, that current techniques cannot analyze. We have proposed TraceSquiz, an online approach to monitor the trace output during endurance test, in order to record only suspicious portions of the trace and discard regular ones. This approach is based on anomaly detection techniques. Our detailed experiments have shown that our approach has a good anomaly detection performance, and can reduce the size of an output trace by an order of magnitude [24]. Serge Emteu successfully defended his PhD about this work on the 15/12/2015 [5].

7.3. Causal reasoning and argumentation

Participants: Philippe Besnard, Louis Bonneau de Beaufort, Marie-Odile Cordier, Yves Moinard.

7.3.1. Searching for explanations from causal relations and ontology for argumentation

We have continued our work on reasoning (precisely search for explanations) from causal relations and ontology [48]. We resort to a well-known model [49] in computational argumentation in order to provide some structure to the collection of potential explanations given by our causal formalism. We have developed a case study, namely the Xynthia storm case, (February 2010, western France, trial September 2014) for which there exists a huge amount of data from various official reports. We have implemented an ASP program which thereby provides another application, besides those already mentioned: mining and landscape simulation, for ASP.

7.3.2. Cognitive maps and Bayesian causal maps

Cognitive map is a qualitative decision model which is frequently used in social science and decision making applications. This model allows to easily organize individuals' judgments, thinking or beliefs about a given problem in a graphical representation containing different concepts and influences between them. However, cognitive maps cannot model uncertainty within the variables and provides only deductive reasoning (predicting an effect given a cause). In [37], we show how to translate the knowledge represented in cognitive maps in the form of arguments and attack relations among them. Given a decision problem, we propose to build, first, a cognitive map by eliciting knowledge from experts and then to transform it into a weighted argumentation framework (WAF for short) for ensuring efficient reasoning. Another contribution concerns enriching the WAF obtained from a given cognitive map for dealing with dynamics through the consideration of a varying set of observations.

Cognitive maps and Bayesian networks are useful formalisms to address knowledge representation. Cognitive maps are powerful graphical models for gathering or displaying knowledge but while offering an easy means to express individuals judgments, drawing inferences remains a difficult task. Bayesian networks are widely used for decision making processes that face uncertain information or diagnosis but are difficult to elicitate. To take advantage of both formalisms and to overcome their drawbacks, Bayesian causal maps (BCM) were developed [75]. In [6], we propose to start from a causal map to construct the model and then set the conditional probabilities. Once the common causal map (CM) is built we can transform it into a BCM which combines causal modeling techniques and bayesian probability theory. We have developed a complete framework and applied it on a real problem in an environmental context. The implemented decision facilitating tool enables the representation of different shellfish dredgers views about their activity as well as the test of different fishery management scenarios.

EX-SITU Team

7. New Results

7.1. Fundamentals of Interaction

Participants: Sarah Fdili Alaoui, Michel Beaudouin-Lafon, Cédric Fleury, Wendy Mackay, Theophanis Tsandilas.

In order to better understand fundamental aspects of interaction, ExSitu studies interaction in extreme situations. We conduct indepth observational studies and controlled experiments which contribute to theories and frameworks that unify our findings and help us generate new, advanced interaction techniques. Although we continue to explore the theory of Instrumental Interaction in the context of multi-surface environments [23], we are also extending it into a wider framework we call *information substrates*. This has resulted in several prototypes, such as Webstrates [18]. We also continue to study elementary interaction tasks in large-scale environments, such as pointing [11] and object manipulation [15].

Information substrates – “Instrumental interaction” argues that, since our interaction with the physical world is often mediated by tools, or instruments, we should do the same in the digital world. Our work on multisurface environments has demonstrated the value of this model, for example, to support distributed interfaces in which the user controls the content of a wall-sized display using handheld devices [23]. Instrumental interaction does not, however, describe the “objects of interest” that instruments interact with, nor does it explain how an object becomes an instrument, nor how users appropriate them in unexpected ways (the principle of “co-adaptation”).

“Information substrates” embrace a wider scope than instrumental interaction: both objects and instruments are “substrates” that hold information and behavior, and can be combined in arbitrary ways. What makes an object an instrument is defined not by what it is but by how the user uses it. We started to explore this concept with Webstrates [23], a web-based environment in which content and tools are embedded in the same information substrate—in this case the Document Object Model (DOM) (Figure 7).

Our work on information substrates has influenced other projects in the group. For example, our work on tools to help programmers parallelize and optimize their code [22] uses coordinated views of the code: a traditional text view and a graphical polyhedral visualization (Figure 2). These two substrates afford different types of manipulation by the user, but share the same underlying information, i.e. the algorithm being designed. The SketchSliders technique [20], described in the following section, provides users with an easily customizable approach to control complex visual displays. SketchSliders act as a substrate for creating slider instruments, which are independent from but tightly coupled with the visual display they control. By letting users define their own sliders, we solve the long-standing problem of combining power and simplicity. Finally, the ColorLab prototypes [17], described in the following section, provide artists and graphic designers with substrates that offer novel ways to interact with and display color relationships.

Interaction in the large – ExSitu and its predecessor InSitu have a long history of studying the most fundamental action in visual environments: pointing. We recently published an extensive 64-page journal article [11] on our studies of pointing on large, wall-sized displays. In such environments, users must be able to point from a distance, typically up to a few meters from the screen, with great accuracy. Existing techniques are ill-suited for this task, due to the combination of the high index of difficulty and the constraint that users must be able to move around in the room while pointing.

We have designed and tested a number of techniques, including dual-mode techniques that combine coarse pointing with direct techniques, such as ray-casting or using the orientation of the head, and fine pointing with relative techniques, such as using a hand-held touchpad 3. Rather than proposing the “ultimate” pointing technique for such environments, we provide a set of criteria, a set of techniques derived from those criteria, and a calibration technique for optimizing the transfer functions used by relative pointing techniques under extreme conditions.

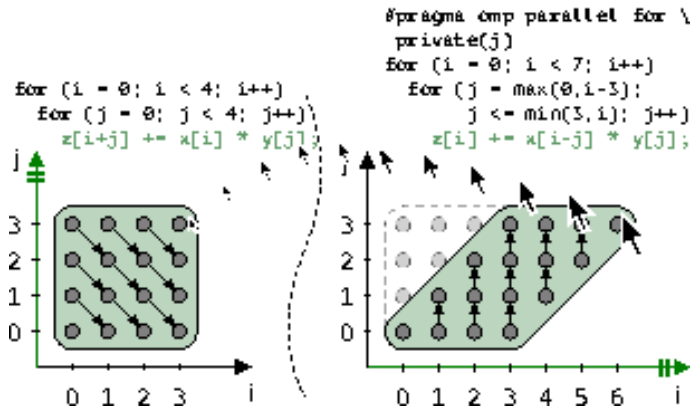


Figure 2. Performing a skew transformation to parallelize polynomial multiplication. The code is automatically transformed from its original form (left) to the skewed one (right).

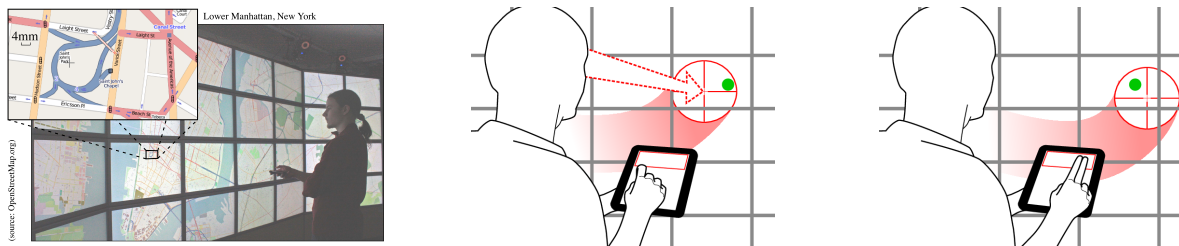


Figure 3. The challenge of pointing on a wall from a distance on a ultra-high resolution wall-sized display (left). Two of the pointing techniques that we evaluated: coarse pointing using the orientation of the head (center) vs. a two-finger swipe (right). In both cases, a one-finger swipe controls precise pointing.

In collaboration with the Inria REVES group in Sophia Antipolis, we proposed a framework for analyzing 3D object manipulation in immersive environments [15]. We decomposed 3D object manipulation into the component movements, taking into account both physical constraints and mechanics. We then fabricated five physical devices that simulate these movements in a measurable way under experimental conditions. We implemented the devices in an immersive environment and conducted an experiment to evaluate direct finger-based against ray-based object manipulation. We identified the compromises required when designing devices that (i) are reproducible in both real and virtual settings, and (ii) can be used in experiments to measure user performance.

7.2. Creativity

Participants: Sarah Fdili Alaoui, Michel Beaudouin-Lafon, Ghita Jalal, Wendy Mackay, Joseph Malloch, Nolwenn Maudet, Theophanis Tsandilas.

ExSitu is interested in understanding the work practices of creative professionals, particularly artists, designers, and scientists, who push the limits of interactive technology. This year, we conducted studies and created tools for a variety of such users. Based on contextual interviews with artists, designers and scientists, we created the *Color Portraits* design space [17] to characterize color manipulation activities, which influenced the design of a set of color manipulation tools (*Color Lab*). We designed *BricoSketch* [21] to enable professional illustrators to work at different levels of detail on paper. We studied how makers *remix* each others' designs by analyzing metadata from over 175,000 digital designs from Thingiverse [19]. We created *SketchSliders* [20] to help scientists explore their data by sketching and manipulating free-form interactive controllers. Finally, we studied the meaning and use of the term *evaluation* within the NIME (New Interfaces for Musical Expression) community [14].

Our studies of these “extreme users” allows us to obtain empirical grounding for the theoretical concepts of instrumental interaction, information substrates and co-adaptive systems. We expect to transfer what we learn to the design of creative tools, first for expert users, then for non-specialists and non-professional users.

Color Portraits – We conducted contextual interviews with 16 participants, who provided detailed examples of how they used color to create 69 different artistic or technical artifacts [17]. Based on results from these interviews, we created the Color Portraits design space to help identify color manipulation requirements that are poorly addressed by today's color manipulation tools. We then developed a set of novel color-manipulation tools that test the generative power of the design space. We presented these to users as probes. Our observations of how users interacted with the color probes provide implications for the design of more advanced tools.

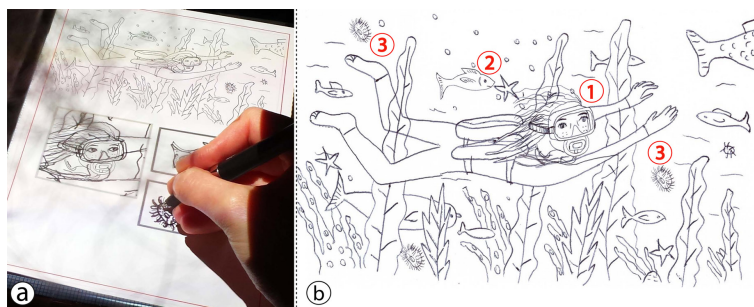


Figure 4. An artist drawing on paper with *BricoSketch*. (a) The artist has created three views on paper to draw parts of the illustration with higher detail: (1) the head of the diver, (2) a fish, and (3) an urchin. (b) The final composition after blending the partial views together.

BricoSketch – We conducted interviews with four professional illustrators and investigated how they use technology and paper in their creative process [21]. We also studied the evolution of the work of one of these illustrators for a period of two years. In interaction with this artist, we designed BricoSketch. BricoSketch enables illustrators to interactively create partial views of their drawings. Such views can be transposed and rescaled. Artists can then use them to create variations of their illustrations or add details with higher drawing precision. Our implementation is based on interactive paper technology that allows for above-the-surface interaction and supports traditional drawing tools such as common pens and pencils.

Remixing Designs – We investigated [19] how makers remix digital designs for physical objects on “Thingiverse”, a well-established online 3D-printing maker community. We collected metadata from over 175,000 digital designs and analyzed the *remixing graph* – links between sources and remixes that primarily exhibit a tree-like inheritance structure. We also used this data to identify particularly influential and surprising “Things”, which we further examined via qualitative case studies. We concluded with specific suggestions for online design repositories and design software so as to provide better support for remixing, and thus build stronger online maker communities.

SketchSliders – We developed SketchSliders [20], range sliders that users can freely sketch on a mobile device to parametrize and customize their data exploration on a wall display. With a small combination of sketches and gestures, users can create complex interactive controllers, such as slider branches and data transformation sliders 5. In addition to their natural custom shape, the sketched sliders can also be enhanced by interaction aids such as slider cursors, markers and distribution visualizations. We evaluated the sketching interface with six visualization experts and found that SketchSliders accommodate a wide range of exploration strategies, as well as help users focus and customize their visual explorations.

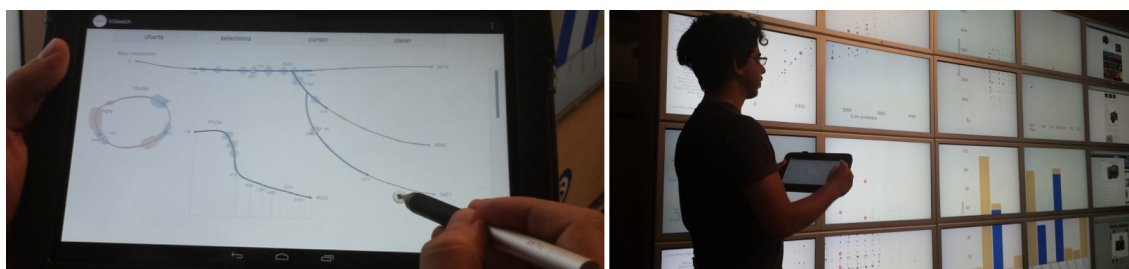


Figure 5. SketchSliders let users directly sketch viusalization controllers to explore multidimensional datasets.

Evaluation for NIME – We explored the use of evaluation techniques and terminology within the past three years of the New Interfaces for Music Expression (NIME) conference [14]. We categorized each paper that mentioned evaluation according to five criteria: a) targets and stakeholders considered, b) goals set, c) criteria used, d) methods used, and e) duration of evaluation. Results suggest that the NIME community does not share a common culture with respect to evaluation, with little consistency regarding use of the term. This paper raises the issue of evaluation within NIME community, with the goal of using it more consistently and effectively in the future.

7.3. Collaboration

Participants: Michel Beaudouin-Lafon, Cédric Fleury, Wendy Mackay, Can Liu, Ignacio Avellino Martinez.

ExSitu is interested in exploring new ways to support collaborative interaction, especially within and across large interactive spaces such as those of the Digiscope network (<http://digiscope.fr/>). We started to investigate how to support telepresence among large, heterogeneous interactive spaces [24], [25]. In particular, we studied how accurately a user can interpret deictic gestures in a video feed of a remote user [12]. These deictic

gestures are important for conveying non-verbal cues for communication between remote users. We also created *Webstrates* [18], an environment for exploring shareable dynamic media and the concept of *information substrate*.

Telepresence among large, heterogeneous interactive spaces – Large interactive spaces are powerful tools that can help scientific, industrial and business users to collaborate on large and complex data sets. In order to reach their full potential, these spaces must not only support local collaboration, but also collaboration with remote users, who may have significantly different display and interaction capabilities, such as a wall-display connected to an immersive CAVE.

We explain why supporting telepresence across large interactive spaces is critical for remote collaboration [24]. We have also started to explore how such asymmetric interaction capabilities provide interesting opportunities for new collaboration strategies in large interactive spaces [25].

Accuracy of deictic gestures for telepresence – In the context of telepresence on large wall-sized displays, we investigated how accurately a user can interpret the video feed of a remote user showing a shared object on the display, by looking at it or by looking and pointing at it (Figure 6) [12]. We also analyzed how sensitive distance and angle errors are to the relative position between the remote viewer and the video feed. We showed that users can accurately determine the target, that eye gaze alone is more accurate than when combined with the hand, and that the relative position between the viewer and the video feed has little effect on accuracy. These findings can inform the design of future telepresence systems for wall-sized displays.

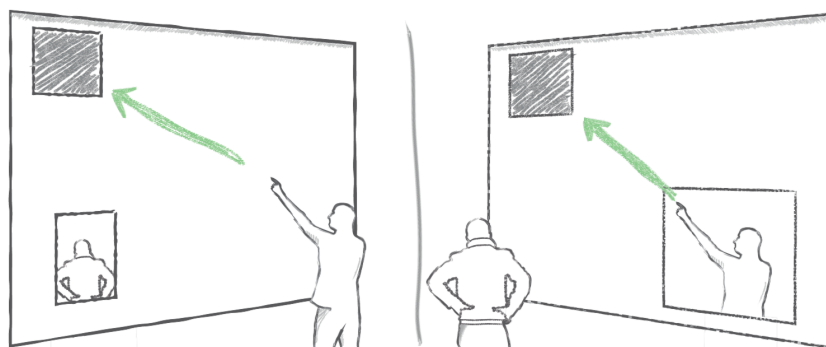


Figure 6. Users working on shared objects using two remote wall-sized displays : a user (left) shows a shared object by pointing at it and the remote user (right) can see which object is being shown through the video feed.

Webstrates – In collaboration with Université of Aarhus (Denmark) and Institut Mines Telecom, we created *Webstrates* [18], a system inspired by Alan Kay's early vision of interactive dynamic media. *Webstrates* is based on web technology: web pages served by the *Webstrates* server can be shared in real time among multiple users, on any web-enabled device. By using transclusion, a *webstrate* page can include other *Webstrates*. *Webstrates* can also include code, making them dynamic and interactive. A *Webstrate* that can act on another, transcluded *Webstrate*, is similar to an editor on a classical desktop environment. However the distinction between content and tools, documents and applications is blurred, e.g. content can be used as a tool, and tools can be shared like regular content. We implemented two case studies to illustrate *Webstrates* (Figure 7). We authored the article collaboratively, using functionally and visually different editors that we could personalize and extend at run-time. We also used *Webstrates* to orchestrate a presentation, using multiple devices to control the presentation, to let the audience participate and the session chair organize the session. We demonstrated the simplicity and generative power of *Webstrates* with three additional prototypes and evaluated them from a systems perspective. *Webstrates* runs in our WildOS middleware on the WILD and WILDER rooms, and is used for some of our projects on telepresence.



Figure 7. Sample uses of Webstrates: (a) Collaborative document authoring with different editors personalized at run-time; (b) Multiple devices used to sketch a figure (tablet 1), see it in a print preview (tablet 2), and adjust it in a graphics editor (laptop). (c) Distributed talk controlled remotely by a speaker with a separate interface for audience participation.

EXMO Project-Team

7. New Results

7.1. Ontology matching and alignments

We pursue our work on ontology matching and alignment support [4] with contributions to evaluation and the use of algebras of relations within alignments.

7.1.1. Evaluation

Participant: Jérôme Euzenat [Correspondent].

Since 2004, we run the Ontology Alignment Evaluation Initiative (OAEI) which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [3].

This year, we also handed out the organisation of OAEI 2015 to Ernesto Jiménez Ruiz (University of Oxford). We used again our generator for generating new version of benchmarks. The Alignment API was used for manipulating alignments and evaluating results [8].

The participating systems and evaluation results were presented in the 10th Ontology Matching workshop [13], held Bethlehem (PA US). More information on OAEI can be found at <http://oaei.ontologymatching.org/>.

7.1.2. Algebras of alignment relations

Participants: Armen Inants [Correspondent], Jérôme Euzenat.

Qualitative calculi are central in qualitative binary constraint satisfaction problems. All formalisms developed so far are homogeneous – they assume a single universe. We had previously shown the advantages of using a homogeneous qualitative calculus for expressing ontology alignment relations between concepts.

They make it possible to aggregate alignments disjunctively or conjunctively and to propagate alignments within a network of ontologies. The previously considered algebra of relations contains taxonomical relations between classes only. We have tackled the problem of combining two or more calculi over disjoint universes into a single calculus [9]. The problem is important because ontology matching deals with various kinds of ontological entities: concepts, individuals, properties. We have designed an algorithm for combining two homogeneous calculi with different universes into a single calculus. This has been applied to alignment relations [9] combining algebras for relations between concepts and individuals. It is, first, able to deal with empty classes, and, second, incorporates all qualitative taxonomical relations that occur between individuals and concepts, including the relations “is a” and “is not”. We have proved that this algebra is coherent with respect to the simple semantics of alignments.

The proposed algebras of relations and others have been integrated within the Alignment API (§6.1).

This work is part of the PhD of Armen Inants.

7.2. Data interlinking

The web of data uses semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together. It is thus important to be able to establish links between these data, both for the web of data and for the semantic web that it contributes to feed. We consider this problem from different perspectives.

7.2.1. Interlinking cross-lingual RDF data sets

Participants: Tatiana Lesnikova [Correspondent], Jérôme David, Jérôme Euzenat.

RDF data sets are being published with labels that may be expressed in different languages. Even systems based on graph structure, ultimately rely on anchors based on language fragments. In this context, data interlinking requires specific approaches in order to tackle cross-lingualism. We proposed a general framework for interlinking RDF data in different languages and implemented two approaches: one approach is based on machine translation, the other one takes advantage of multilingual references, such as BabelNet. This year we investigated the second approach [10], finding that results were not as good as the translation approach. We also conducted evaluations on TheSoz, Agrovoc and Eurovoc thesauri.

This work is part of the PhD of Tatiana Lesnikova developed in the LINDICLE project (§9.1.1).

7.2.2. *An iterative import-by-query approach to data interlinking*

Participant: Manuel Atencia Arcas [Correspondent].

We modelled the problem of data interlinking as a reasoning problem on possibly decentralised data. We described an import-by-query algorithm that alternates steps of sub-query rewriting and of tailored querying of data sources [11]. It only imports data as specific as possible for inferring or contradicting target owl:sameAs assertions. Experiments conducted on a real-world dataset have demonstrated in practice the feasibility and usefulness of this approach for data interlinking and disambiguation purposes.

Additionally, and in line with the problem of dealing with uncertainty in linked data, we have proposed a probabilistic mechanism of trust that allow peers in a semantic peer-to-peer network to select the peers that are better suited to answer their queries, when query reformulation based on alignments may be unsatisfactory due to unsoundness or incompleteness of alignments [5].

This work was carried out in collaboration with Mustafa Al-Bakri and Marie-Christine Rousset (LIG).

7.2.3. *Link key extraction*

Participants: Jérôme David [Correspondent], Manuel Atencia Arcas, Jérôme Euzenat.

Ontologies do not necessarily come with key descriptions, and never with link key assertions (§3.3). Keys can be extracted from data by assuming that keys holding for specific data sets, may hold universally.

Following the work of last year on link key extraction [1] and the characterisation of the approach in formal concept analysis, we have fully characterised the results of our algorithm as formal concepts. We have also plans for extending both the approach and its formal concept analysis description through (i) applying it to full link keys as described in §3.3, (ii) applying it to join and hierarchical key extraction, and (iii) applying it to hierarchical key extraction.

This work has been developed partly in the LINDICLE project (§9.1.1). Formal concept analysis aspects are considered with Amedeo Napoli (Orpailleur, LORIA).

7.3. **Dynamic aspects of networks of ontologies**

Huge quantities of data described by ontologies and linked together are made available. These are generated in an independent manner by autonomous providers such as individuals or companies. They are heterogeneous and their joint exploitation requires connecting them.

However, data and knowledge have to evolve facing changes in what they represent, changes in the context in which they are used and connections to new data and knowledge sources. As their production and exchange are growing larger and more connected, their evolution is not anymore compatible with manual curation and maintenance. We work towards their continuous evolution as it is critical to their sustainability.

Two different approaches are currently explored.

7.3.1. *Evolution of ontology networks and linked data*

Participants: Adam Sanchez Ayte [Correspondent], Jérôme David, Jérôme Euzenat.

We are considering the global evolution of knowledge represented by interdependent ontologies, data, alignments and links. Our goal is to be able to maintain such a structure with respect to the processes which are involved in its construction: logical inference, ontology matching, link key extraction, link generation, etc.

Our initial work is focused on how data and ontology changes cause alignment evolution, in particular when the alignment have been produced through instance-based matching using links between data. In this regard, we are developing techniques for circumscribing the elements and relationships affected by the change as well as evaluating the need for change propagation, i.e, most of the time a simple change will not trigger link key recomputation (§7.2.3).

This work is part of the PhD thesis of Adam Sanchez Ayte developed in the LINDICLE project (§9.1.1).

7.3.2. Revision in networks of ontologies

Participant: Jérôme Euzenat [Correspondent].

We reconsidered the belief revision problem in the context of networks of ontologies (§3.2): given a set of ontologies connected by alignments, how to evolve them such that they account for new information. In networks of ontologies, inconsistency may come from two different sources: local inconsistency in a particular ontology or alignment, and global inconsistency between them. Belief revision is well-defined for dealing with ontologies; we have investigated how it can apply to networks of ontologies. We formulated revision postulates for alignments and networks of ontologies based on an abstraction of existing semantics of networks of ontologies. We showed that revision operators cannot be simply based on local revision operators on both ontologies and alignments. We adapted the partial meet revision framework to networks of ontologies and show that it indeed satisfies the revision postulates [7]. Finally, we considered strategies based on network characteristics for designing concrete revision operators.

FLOWERS Project-Team

7. New Results

7.1. Robotic And Computational Models Of Human Development and Cognition

7.1.1. Computational Models Of Information-Seeking, Curiosity And Attention in Humans and Animals

Participants: Manuel Lopes, Pierre-Yves Oudeyer [correspondant], Jacqueline Gottlieb, Adrien Baranes, William Schueller, Sebastien Forestier, Nabil Daddaouda, Nicholas Foley.

This project involves a collaboration between the Flowers team and the Cognitive Neuroscience Lab of J. Gottlieb at Columbia Univ. (NY, US) on the understanding and modeling of mechanisms of curiosity and attention that until now have been little explored in neuroscience, computer science and cognitive robotics. It is organized around the study of the hypothesis that information gain could generate intrinsic reward in the brain (living or artificial), controlling attention and exploration independently from material rewards. The project combines expertise about attention and exploration in the brain and a strong methodological framework for conducting experimentations with monkeys and humans (Gottlieb's lab) together with cognitive modeling of curiosity and learning in the Flowers team.

Such a collaboration paves the way towards a central objective, which is now a central strategic objective of the Flowers team: designing and conducting experiments in animals and humans informed by computational/mathematical theories of information seeking, and allowing to test the predictions of these computational theories.

7.1.1.1. Context

Curiosity can be understood as a family of mechanisms that evolved to allow agents to maximize their knowledge of the useful properties of the world - i.e., the regularities that exist in the world - using active, targeted investigations. In other words, we view curiosity as a decision process that maximizes learning (rather than minimizing uncertainty) and assigns value ("interest") to competing tasks based on their epistemic qualities - i.e., their estimated potential allow discovery and learning about the structure of the world.

Because a curiosity-based system acts in conditions of extreme uncertainty (when the distributions of events may be entirely unknown) there is in general no optimal solution to the question of which exploratory action to take [105], [122], [130]. Therefore we hypothesize that, rather than using a single optimization process as it has been the case in most previous theoretical work [90], curiosity is comprised of a family of mechanisms that include simple heuristics related to novelty/surprise and measures of learning progress over longer time scales [16] [74], [112]. These different components are related to the subject's epistemic state (knowledge and beliefs) and may be integrated with fluctuating weights that vary according to the task context. We will quantitatively characterize this dynamic, multi-dimensional system in the framework of Bayesian Reinforcement Learning, as described below.

Because of its reliance on epistemic currencies, curiosity is also very likely to be sensitive to individual differences in personality and cognitive functions. Humans show well-documented individual differences in curiosity and exploratory drives [103], [129], and rats show individual variation in learning styles and novelty seeking behaviors [88], but the basis of these differences is not understood. We postulate that an important component of this variation is related to differences in working memory capacity and executive control which, by affecting the encoding and retention of information, will impact the individual's assessment of learning, novelty and surprise and ultimately, the value they place on these factors [127], [136], [70], [140]. To start understanding these relationships, about which nothing is known, we will search for correlations between curiosity and measures of working memory and executive control in the population of children we test in our tasks, analyzed from the point of view of a computational model based on Bayesian reinforcement learning.

A final premise guiding our research is that essential elements of curiosity are shared by humans and non-human primates. Human beings have a superior capacity for abstract reasoning and building causal models, which is a prerequisite for sophisticated forms of curiosity such as scientific research. However, if the task is adequately simplified, essential elements of curiosity are also found in monkeys [103], [100] and, with adequate characterization, this species can become a useful model system for understanding the neurophysiological mechanisms.

7.1.1.2. Objectives

Our studies have several highly innovative aspects, both with respect to curiosity and to the traditional research field of each member team.

- Linking curiosity with quantitative theories of learning and decision making: While existing investigations examined curiosity in qualitative, descriptive terms, here we propose a novel approach that integrates quantitative behavioral and neuronal measures with computationally defined theories of Bayesian Reinforcement Learning and decision making.
- Linking curiosity in children and monkeys: While existing investigations examined curiosity in humans, here we propose a novel line of research that coordinates its study in humans and non-human primates. This will address key open questions about differences in curiosity between species, and allow access to its cellular mechanisms.
- Neurophysiology of intrinsic motivation: Whereas virtually all the animal studies of learning and decision making focus on operant tasks (where behavior is shaped by experimenter-determined primary rewards) our studies are among the very first to examine behaviors that are intrinsically motivated by the animals' own learning, beliefs or expectations.
- Neurophysiology of learning and attention: While multiple experiments have explored the single-neuron basis of visual attention in monkeys, all of these studies focused on vision and eye movement control. Our studies are the first to examine the links between attention and learning, which are recognized in psychophysical studies but have been neglected in physiological investigations.
- Computer science: biological basis for artificial exploration: While computer science has proposed and tested many algorithms that can guide intrinsically motivated exploration, our studies are the first to test the biological plausibility of these algorithms.
- Developmental psychology: linking curiosity with development: While it has long been appreciated that children learn selectively from some sources but not others, there has been no systematic investigation of the factors that engender curiosity, or how they depend on cognitive traits.

7.1.1.3. Current results

During the first period of the associated team (2013-2015), we laid the operational foundations of the collaboration resulting in several milestone joint journal articles [110], [90], [84][27], new experimental paradigms for the study of curiosity, and organized a major scientific event: the first international interdisciplinary symposium on information seeking, curiosity and attention (web: <https://openlab-flowers.inria.fr/t/first-interdisciplinary-symposium-on-information-seeking-curiosity-and-attention/21>).

In particular, new results in 2015 include:

7.1.1.4. Eye movements reveal epistemic curiosity in human observers

Saccadic (rapid) eye movements are primary means by which humans and non-human primates sample visual information. However, while saccadic decisions are intensively investigated in instrumental contexts where saccades guide subsequent actions, it is largely unknown how they may be influenced by curiosity – the intrinsic desire to learn. While saccades are sensitive to visual novelty and visual surprise, no study has examined their relation to epistemic curiosity – interest in symbolic, semantic information. To investigate this question, we tracked the eye movements of human observers while they read trivia questions and, after a brief delay, were visually given the answer. We showed that higher curiosity was associated with earlier anticipatory orienting of gaze toward the answer location without changes in other metrics of saccades or fixations, and that these influences were distinct from those produced by variations in confidence and surprise. Across

subjects, the enhancement of anticipatory gaze was correlated with measures of trait curiosity from personality questionnaires. Finally, a machine learning algorithm could predict curiosity in a cross-subject manner, relying primarily on statistical features of the gaze position before the answer onset and independently of covariations in confidence or surprise, suggesting potential practical applications for educational technologies, recommender systems and research in cognitive sciences. We published these results in [27], providing full access to the annotated database allowing readers to reproduce the results. Epistemic curiosity produces specific effects on oculomotor anticipation that can be used to read out curiosity states.

7.1.1.5. *Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates*

Intelligent animals have a high degree of curiosity – the intrinsic desire to know – but the mechanisms of curiosity are poorly understood. A key open question pertains to the internal valuation systems that drive curiosity. What are the cognitive and emotional factors that motivate animals to seek information when this is not reinforced by instrumental rewards? Using a novel oculomotor paradigm, combined with reinforcement learning (RL) simulations, we show that monkeys are intrinsically motivated to search for and look at reward-predictive cues, and that their intrinsic motivation is shaped by a desire to reduce uncertainty, a desire to obtain conditioned reinforcement from positive cues, and individual variations in decision strategy and the cognitive costs of acquiring information. The results suggest that free-viewing oculomotor behavior reveals cognitive and emotional factors underlying the curiosity driven sampling of information. [84]

7.1.2. *Computational Models Of Speech Development: the Roles of Active Learning, Curiosity and Self-Organization*

Participants: Pierre-Yves Oudeyer [correspondant], Clément Moulin-Frier, Sébastien Forestier.

7.1.2.1. *Special issue on the cognitive nature of speech sounds*

Together with Jean-Luc Schwartz and Kenneth de Jong, Flowers members Clément Moulin-Frier and Pierre-Yves Oudeyer guest-edited a milestone special issue of the Journal of Phonetics focusing on theories of the cognitive nature of speech sounds, and with a special emphasis on presenting and analyzing a rich series of computational models of speech evolution and acquisition developed in the past years internationally, including models developed by the guest-editors. The editorial of this special issue was published in [35] and the special issue is accessible at: <http://www.sciencedirect.com/science/journal/00954470/53>.

7.1.2.2. *The COSMO model: A Bayesian modeling framework for studying speech communication and the emergence of phonological systems*

(Note: this model was developed while C. Moulin-Frier was at GIPSA Lab, and writing was partly achieved while he was at Inria). While the origin of language remains a somewhat mysterious process, understanding how human language takes specific forms appears to be accessible by the experimental method. Languages, despite their wide variety, display obvious regularities. In this paper, we attempt to derive some properties of phonological systems (the sound systems for human languages) from speech communication principles. The article [33] introduces a model of the cognitive architecture of a communicating agent, called COSMO (for “Communicating about Objects using Sensory–Motor Operations”) that allows a probabilistic expression of the main theoretical trends found in the speech production and perception literature. This enables a computational comparison of these theoretical trends, which helps us to identify the conditions that favor the emergence of linguistic codes. It presents realistic simulations of phonological system emergence showing that COSMO is able to predict the main regularities in vowel, stop consonant and syllable systems in human languages.

7.1.2.3. *The role of self-organization, motivation and curiosity in speech development and evolution*

In the article [34], Oudeyer discusses open scientific challenges for understanding development and evolution of speech forms. Based on the analysis of mathematical models of the origins of speech forms, with a focus on their assumptions, the article studies the fundamental question of how speech can be formed out of non-speech, at both developmental and evolutionary scales. In particular, it emphasizes the importance of embodied self-organization, as well as the role of mechanisms of motivation and active curiosity-driven exploration in speech formation. Finally, it discusses an evolutionary-developmental perspective of the origins of speech.

7.1.2.4. Robotic models of the joint development of speech and tool use

A scientific challenge in developmental and social robotics is to model how autonomous organisms can develop and learn open repertoires of skills in high-dimensional sensorimotor spaces, given limited resources of time and energy. This challenge is important both from the fundamental and application perspectives. First, recent work in robotic modeling of development has shown that it could make decisive contributions to improve our understanding of development in human children, within cognitive sciences [90]. Second, these models are key for enabling future robots to learn new skills through lifelong natural interaction with human users, for example in assistive robotics [124].

In recent years, two strands of work have shown significant advances in the scientific community. On the one hand, algorithmic models of active learning and imitation learning combined with adequately designed properties of robotic bodies have allowed robots to learn how to control an initially unknown high-dimensional body (for example locomotion with a soft material body [73]). On the other hand, other algorithmic models have shown how several social learning mechanisms could allow robots to acquire elements of speech and language [79], allowing them to interact with humans. Yet, these two strands of models have so far mostly remained disconnected, where models of sensorimotor learning were too “low-level” to reach capabilities for language, and models of language acquisition assumed strong language specific machinery limiting their flexibility. Preliminary work has been showing that strong connections are underlying mechanisms of hierarchical sensorimotor learning, artificial curiosity, and language acquisition [125].

Recent robotic modeling work in this direction has shown how mechanisms of active curiosity-driven learning could progressively self-organize developmental stages of increasing complexity in vocal skills sharing many properties with the vocal development of infants [113]. Interestingly, these mechanisms were shown to be exactly the same as those that can allow a robot to discover other parts of its body, and how to interact with external physical objects [120].

In such current models, the vocal agents do not associate sounds to meaning, and do not link vocal production to other forms of action. In other models of language acquisition, one assumes that vocal production is mastered, and hand code the meta-knowledge that sounds should be associated to referents or actions [79]. But understanding what kind of algorithmic mechanisms can explain the smooth transition between the learning of vocal sound production and their use as tools to affect the world is still largely an open question.

The goal of this work is to elaborate and study computational models of curiosity-driven learning that allow flexible learning of skill hierarchies, in particular for learning how to use tools and how to engage in social interaction, following those presented in [120], [73], [118], [113]. The aim is to make steps towards addressing the fundamental question of how speech communication is acquired through embodied interaction, and how it is linked to tool discovery and learning.

A first question that we study in this work is the type of mechanisms that could be used for hierarchical skill learning allowing to manage new task spaces and new action spaces, where the action and task spaces initially given to the robot are continuous and high-dimensional and can be encapsulated as primitive actions to affect newly learnt task spaces.

We presented preliminary results on that question in a poster session [89] of the ICDL/Epirob conference in Providence, RI, USA in August 2015. In this work, we rely more specifically on the R-IAC and SAGG-RIAC series of architectures developed in the Flowers team and we develop different ways to extend those architectures to the learning of several task spaces that can be explored in a hierarchical manner. We describe an interactive task to evaluate different hierarchical learning mechanisms, where a robot has to explore its motor space in order to push an object to different locations. The task can be decomposed into two subtasks where the robot can first explore how to make movements with its hand and then integrate this skill to explore the task of pushing an object.

In the Simplest First strategy, the agent explores successively but with a fixed curriculum the different tasks to learn in the good order: from the simplest one (learning hand movements given motor parameters) to the more complex one (pushing a block with hand movements) that need knowledge about the simpler task to be learned.

In the Top-Down Guidance strategy, the module learning the more complex task (pushing a block with hand movements) gives goals (hand movements) to be reached by the lower-level module (learning hand movements given motor parameters) that will explore for a while to reach that goal before switching to a new given goal.

We also compare our architectures to the control ones where the robot learns directly the not decomposed task, with a competence-based intrinsic motivation (goal babbling) or a fully random motor babbling.

The results show a better exploration for the Top-Down Guidance than the Simplest First hierarchical exploration strategy, and that learning intermediate representations is beneficial in this setup.

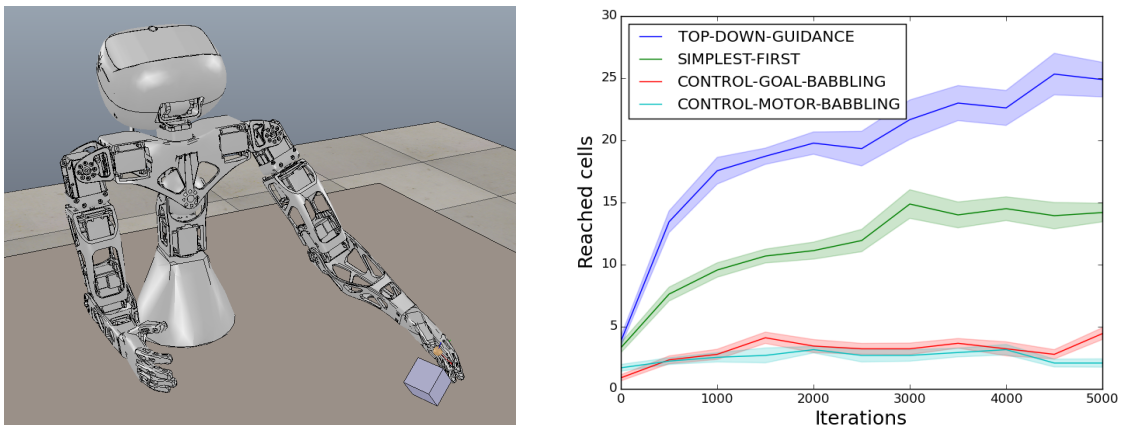


Figure 8. Left: Poppy Torso in the V-REP simulator pushing a block. Right: Exploration results of the different strategies.

7.1.3. Learning in Adult-Child and Human-Robot Interaction

Participants: Anna-Lisa Vollmer [correspondant], Pierre-Yves Oudeyer.

Learning in Adult-Child and Human-Robot Interaction

7.1.3.1. The Change of 'Motionese' Parameters Depending on Children's Age.

Two adult-child interaction studies were analyzed with the focus on the parental teaching behavior, in particular on *motionese* parameters (modifications of child-directed movement). In the first cross-sectional study, parental action demonstrations to three groups of 8–11, 12–23 and 24–30 month-olds ($N = 84$ parents) were investigated. The youngest group of participants was investigated longitudinally in the second study ($N = 18$ parents). Together the results suggest that some motionese parameters (motion pauses, velocity, acceleration) persist over different ages while other parameters (action length, roundness and pace) occur predominantly in the younger group and seem to be primarily used to attract infants' attention on the basis of movement. In contrast, parameters appearing to be more in charge of structuring the action by organizing it in motion pauses seem to persist. We discuss the results in terms of facilitative vs. pedagogical learning in a paper currently under review for the Journal of Experimental Child Psychology.

7.1.3.2. An Alternative to Mapping a Word onto a Concept in Language Acquisition: Pragmatic Frames

According to the mapping metaphor, for a child to learn a word, s/he has to map a word onto a concept of an object/event. We are not convinced that associations can explain word learning, because even though children's attention is on the objects, they do not necessarily remember the connection of the word with the referent. In this theoretical paper, we propose an alternative to the mapping process that is classically assumed as a mechanism for word learning. Our main point holds that word learning is a task, in which children accomplish a goal in cooperation with a partner. In our approach, we follow Bruner's (1983) idea and further

specify pragmatic frames as learning units that drive language acquisition and cognitive development. These units consist of a sequence of language and actions that are co-constructed with a partner to achieve a joint goal. We elaborate on this alternative, offer some initial parametrizations of the concept and embed it in the current language learning approaches in a paper currently under review for *Frontiers in Psychology*, section Cognitive Science.

7.1.3.3. Meta-Analysis of Pragmatic Frames in Human-Robot Interaction for Learning and Teaching: State-of-the-Art and Perspectives

One of the big challenges in robotics today is to learn from inexperienced human users. Despite tremendous research efforts and advances in human-robot interaction (HRI) and robot learning in the past decades, learning interactions with robots remain brittle and rigidly organized, and often are limited to learning only one single task. In this work, we applied the concept of pragmatic frames known from developmental research in humans in a meta-analysis of current approaches on robot learning. This concept offers a new research perspective in HRI as multiple flexible interaction protocols can be used and learned to teach/learn multiple kinds of skills in long-term recurring social interaction. This perspective, thus, emphasizes teaching as a collaborative achievement of teacher and learner. Our meta-analysis focuses on robot learning from a human teacher with respect to the pragmatic frames they (implicitly) use. We show that while the current approaches offer a variety of different learning and teaching behavior, they all employ highly pre-structured, hard-coded pragmatic frames. Compared to natural human-human interaction, interactions are lacking flexibility and expressiveness, and mostly are hardly viable for being realized with truly naive and uninstructed users. We elaborated an outlook on the future research direction with its relevant key challenges that need to be solved for leveraging pragmatic frames for robot learning. These results have been submitted to the *Frontiers in Neurobotics Journal*.

7.1.3.4. Alignment to the Actions of a Robot

Alignment is a phenomenon observed in human conversation: Dialog partners' behavior converges in many respects. Such alignment has been proposed to be automatic and the basis for communicating successfully. Recent research on human-computer dialog promotes a mediated communicative design account of alignment according to which the extent of alignment is influenced by interlocutors' beliefs about each other. Our work aims at adding to these findings in two ways. a) Our work investigates alignment of manual actions, instead of lexical choice. b) Participants interact with the iCub humanoid robot, instead of an artificial computer dialog system. Our results confirm that alignment also takes place in the domain of actions. We were not able to replicate the results of the original study in general in this setting, but in accordance with its findings, participants with a high questionnaire score for emotional stability and participants who are familiar with robots align their actions more to a robot they believe to be basic than to one they believe to be advanced. Regarding alignment over the course of an interaction, the extent of alignment seems to remain constant, when participants believe the robot to be advanced, but it increases over time, when participants believe the robot to be a basic version. These results were published in [38].

7.1.4. Models of Multimodal Concept Acquisition with Non-Negative Matrix Factorization

Participants: Pierre-Yves Oudeyer, Olivier Mangin [correspondant], David Filliat, Louis Ten Bosch.

In the article [32] we introduced MCA-NMF, a computational model of the acquisition of multi-modal concepts by an agent grounded in its environment. More precisely our model finds patterns in multimodal sensor input that characterize associations across modalities (speech utterances, images and motion). We propose this computational model as an answer to the question of how some class of concepts can be learnt. In addition, the model provides a way of defining such a class of plausibly learnable concepts. We detail why the multimodal nature of perception is essential to reduce the ambiguity of learnt concepts as well as to communicate about them through speech. We then present a set of experiments that demonstrate the learning of such concepts from real non-symbolic data consisting of speech sounds, images, and motions. Finally we consider structure in perceptual signals and demonstrate that a detailed knowledge of this structure, named compositional understanding can emerge from, instead of being a prerequisite of, global understanding. An open-source implementation of the MCA-NMF learner as well as scripts and associated experimental data to reproduce the experiments are publicly available.

The python code and datasets allowing to reproduce these experiments and results are available at: <https://github.com/omangin/multimodal>.

7.1.5. Models of Self-organization of lexical conventions: the role of Active Learning in Naming Games

Participants: William Schueller [correspondant], Pierre-Yves Oudeyer.

Our work focuses on the Naming Games framework [135], meant to simulate lexicon evolution in a population from interactions at the individual level. A quite diverse subset of the possible scenarios and algorithms has already been studied, and those do lead to the self-organization of a shared lexicon (understood as associations between meanings and words). However, high values for some parameters (population size, number of possible words and/or meanings that can be referred to) can lead to really slow dynamics. Following the introductory work done in [119], we introduced a new measure of vocabulary evolution based on information theory, as well as various active learning mechanisms in the Naming Games framework allowing the agents to choose what they talk about according to their past. We showed that it improves convergence dynamics in the studied scenarios and parameter ranges. Active learning mechanisms use the confidence an agent has on its own vocabulary (is it already widely used in the population or not?) to choose between exploring new associations (growing vocabulary) or strengthening already existing ones (spreading its own associations to other agents). This was presented at the ICDL/Epirob conference in Providence, RI, USA in August 2015 [59].

A follow-up to this work consisted of changing slightly the base algorithms, allowing agents to select what they want the others to talk about instead of selecting what they would talk about (hearer's choice scenario, the original one being speaker's choice scenario). In the class of algorithms used, with active learning, it leads to faster convergence, with increased robustness to change in parameter values.

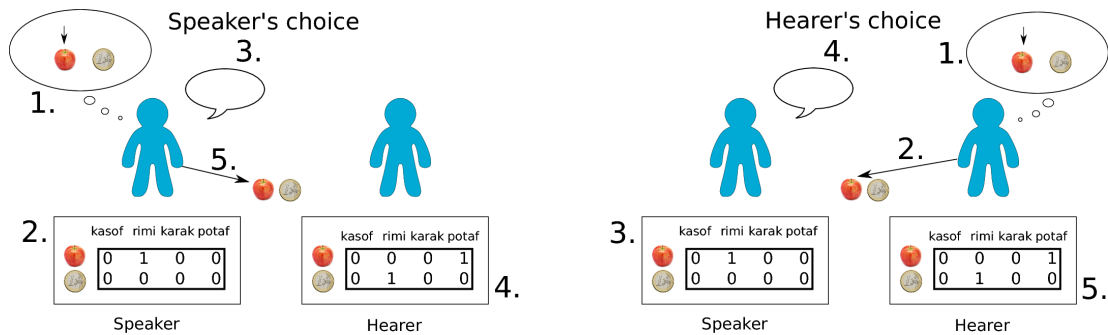


Figure 9. Interaction processes in both active scenarios considered in our work. Beforehand, two individuals have been randomly selected among a population, designated as speaker (S) and hearer (H). Speaker's choice: 1. S chooses a topic, 2. S checks its vocabulary to find or invent an associated word, 3. S utters the word, 4. H guesses the intended meaning, 5. S indicates the intended meaning. Hearer's choice: 1. H chooses a topic, 2. H indicates the intended meaning, 3. S checks its vocabulary to find or invent an associated word, 4. S utters the word, 5. H checks its vocabulary for a meaning associated to the uttered word. In both cases, if all meanings match, the interaction is considered a success, otherwise a failure. After the process, both agents can update their vocabularies to take the interaction into account.

All the simulations can be easily rerun using the provided code and explanatory notebooks on <https://github.com/flowersteam/naminggamesal>.

7.2. Life-Long Robot Learning And Development Of Motor And Social Skills

7.2.1. Uncalibrated BCI

Participants: Manuel Lopes [correspondant], Pierre-Yves Oudeyer, Jonathan Grizou, Inaki Iturrate, Luis Montesano.

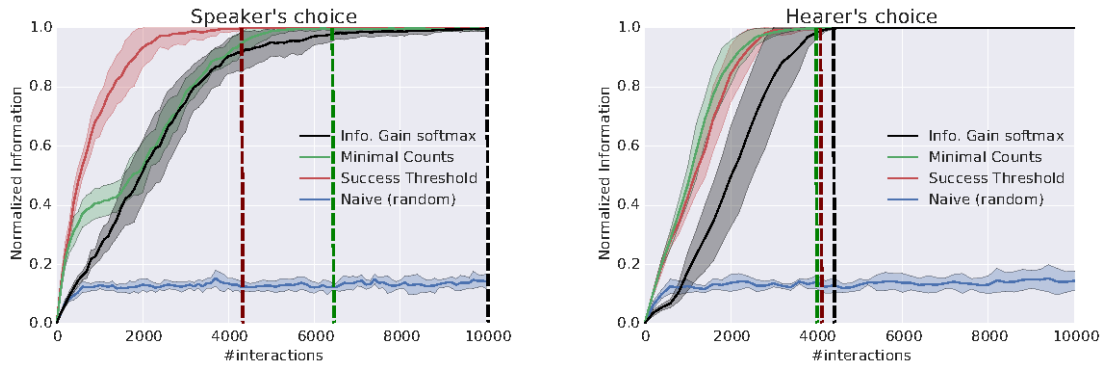


Figure 10. Strategy comparisons in both active scenarios. The measure indicates level of convergence towards a population-wide shared vocabulary (a value of 1 means every agent has the exact same vocabulary). Naive (random, no active learning) strategy converges slowly. Hearer's choice policy is more efficient for all active learning strategies. Last 5% of information are acquired slower when the speaker is choosing. Vertical lines show full convergence time for each strategy. (number of meanings = number of words = population size = 20, averaged over 8 trials)

We developed an new approach for self-calibration BCI for reaching tasks using error-related potentials. The proposed method exploits task constraints to simultaneously calibrate the decoder and control the device, by using a robust likelihood function and an ad-hoc planner to cope with the large uncertainty resulting from the unknown task and decoder. The method has been evaluated in closed-loop online experiments with 8 users using a previously proposed BCI protocol for reaching tasks over a grid. The results show that it is possible to have a usable BCI control from the beginning of the experiment without any prior calibration. Furthermore, comparisons with simulations and previous results obtained using standard calibration hint that both the quality of recorded signals and the performance of the system were comparable to those obtained with a standard calibration approach. [30]

7.2.2. Learning from Demonstration

Participants: Manuel Lopes, Thibaut Munzer [correspondant], Marc Toussaint, Li Wang Wu, Yoan Mollard, Andrea Baisero, Bilal Piot, Matthieu Geist, Olivier Pietquin.

Learning from Demonstration

7.2.2.1. Relational Activity Processes for Modeling Concurrent Cooperation

In multi-agent domains, human-robot collaboration domains, or single-robot manipulation with multiple end-effectors, the activities of the involved parties are naturally concurrent. Such domains are also naturally relational as they involve multiple objects, multiple agents, and models should generalize over objects and agents. We propose a novel formalization of relational concurrent activity processes that allows us to transfer methods from standard (relational) MDPs, such as Monte-Carlo planning and learning from demonstration, to concurrent cooperation domains. We formally compare the formulation to previous propositional models of concurrent decision making and demonstrate the planning and learning from demonstration methods on a real-world human-robot assembly task.

7.2.2.2. Interactive Learning

In paper [56] we consider that robot programming can be made more efficient, precise and intuitive if we leverage the advantages of complementary approaches such as learning from demonstration, learning from feedback and knowledge transfer. We designed a system that, starting from low-level demonstrations of assembly tasks, is able to extract a high-level relational plan of the task. A graphical user interface (GUI)

allows then the user to iteratively correct the acquired knowledge by refining high-level plans, and low-level geometrical knowledge of the task. A final process allows to reuse high-level task knowledge for similar tasks in a transfer learning fashion. We conducted a user study with 14 participants asked to program assembly tasks of small furniture (chair and bench) to validate this approach. The results showed that this combination of approaches leads to a faster programming phase, more precise than just demonstrations, and more intuitive than just through a GUI.

7.2.2.3. Inverse Reinforcement Learning in Relational Domains

We introduced a first approach to the Inverse Reinforcement Learning (IRL) problem in relational domains. IRL has been used to recover a more compact representation of the expert policy leading to better generalize among different contexts. Relational learning allows one to represent problems with a varying number of objects (potentially infinite), thus providing more generalizable representations of problems and skills. We show how these different formalisms can be combined by modifying an IRL algorithm (Cascaded Supervised IRL) such that it handles relational domains. Our results indicate that we can recover rewards from expert data using only partial knowledge about the dynamics of the environment. We evaluate our algorithm in several tasks and study the impact of several experimental conditions such as: the number of demonstrations, knowledge about the dynamics, transfer among varying dimensions of a problem, and changing dynamics. This was published in [49]

7.2.3. A Unified Model for Regression

Regression is the process of learning relationships between inputs and continuous outputs from example data, which enables predictions for novel inputs. Regression lies at the heart of imitation learning, and value function approximation for reinforcement learning. In [37], we provide a novel perspective on regression, by distinguishing rigorously between the models and representations assumed in regression, and the algorithms used to train the parameters of these models. A rather surprising insight is that many regression algorithms⁰ use very similar models; in fact, we show that the algorithm-specific models are *all* special cases of a “unified model”. This perspective clearly separates between representations and algorithms, and allows for a modular exchange between them, for instance in the context of evolutionary optimization.

7.2.4. Multiple Virtual Guides

In co-manipulation, humans and robots solve manipulation tasks together. Virtual guides are important tools for co-manipulation, as they constrain the movement of the robot to avoid undesirable effects, such as collisions with the environment. Defining virtual guides is often a laborious task requiring expert knowledge. This restricts the usefulness of virtual guides in environments where new tasks may need to be solved, or where multiple tasks need to be solved sequentially, but in an unknown order.

To this end, we have proposed a framework for *multiple probabilistic virtual guides*, and demonstrated a concrete implementation of such guides using kinesthetic teaching and Gaussian mixture models [57], [58]. Our approach enables non-expert users to design virtual guides through demonstration. Also, they may demonstrate novel guides, even if already known guides are active. Finally, users are able to intuitively select the appropriate guide from a set of guides through physical interaction with the robot.

7.2.5. Legible Motion

Participants: Manuel Lopes, Baptiste Busch [correspondant], Jonathan Grizou, Freek Stulp.

⁰Locally Weighted Regression, Receptive Field Weighted Regression, Locally Weighted Projection Regression, Gaussian Mixture Regression, Model Trees, Radial Basis Function Networks, Kernel Ridge Regression, Gaussian Process Regression, Support Vector Regression, Random Features Regularized Least Squares, Sparse Spectrum Gaussian Process Regr., Regression Trees, Extreme Learning Machines.

In a human-robot collaboration context, understanding and anticipating the robot intentions ease the completion of a joint-task. Whereas previous work has sought to explicitly optimize the legibility of behavior, we investigate legibility as a property that arises automatically from general requirements on the efficiency and robustness of joint human-robot task completion. We propose an optimization algorithm, based on policy improvement, that brings out the most legible robot's trajectories during the interaction (cf. Figure 11). The conducted user study highlights that humans become better at predicting sooner the robot's intentions. This leads to faster and more robust overall task completion. This work have been published to IROS 2015[60] and was submitted to the International Journal of Social Robotics under the special issue: Towards a framework for Joint Action.

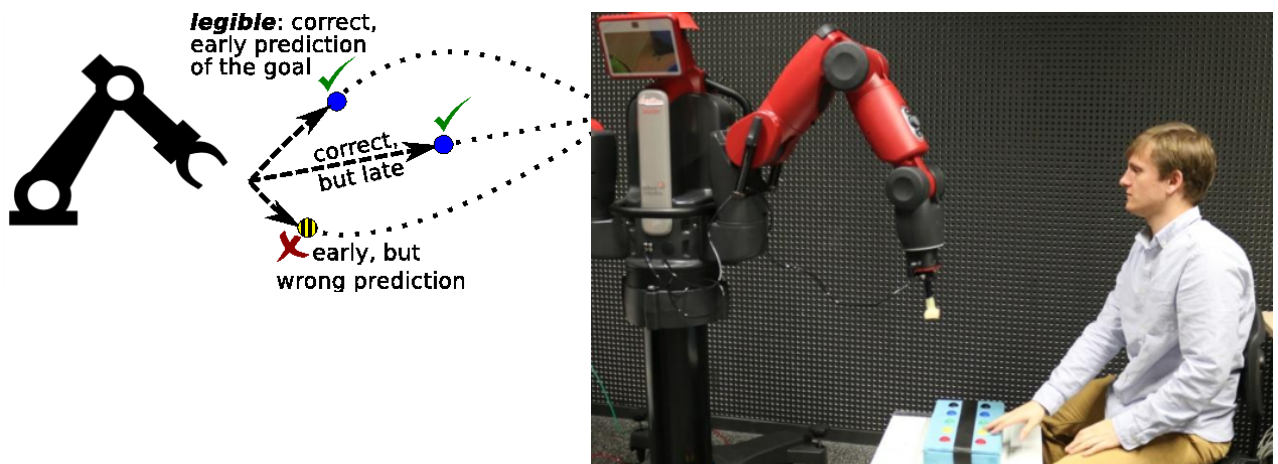


Figure 11. Illustration of the experimental setup. The robots aims for a button and press it. The human predict the target and is instructed to quickly press the button of the same color when sufficiently confident about this prediction.

7.2.6. Demonstrator of human-robot interface for teaching a collaborative task in the context of assistive robotics

Participants: Pierre Rouanet [correspondant], Yoan Mollard, Thibaut Munzer, Baptiste Busch, Manuel Lopes, Pierre-Yves Oudeyer.

In the context of the Roméo 2 project, we have developed a demonstrator of a human-robot interface designed for non-expert users. It allows them to teach a new collaborative task to a robot through simple and intuitive interactions. It is based on the approach of inverse reinforcement learning in relational domains described above.

The context of the demonstrator is assistive robotics where typically an elderly person wants to teach a robot (we use Baxter in this case) how it can help him to prepare a meal. For instance, the user will show the robot that first he wants the robot to hold its bowl and that he stirs it. Then, the robot should put the bowl on a plate. Then, the user will teach the robot that he wants the robot to grab a glass and put it on the right of the bowl...

7.2.7. Diversity-driven curiosity-driven learning and transfer learning

Participants: Fabien Benureau [correspondant], Pierre-Yves Oudeyer.

7.2.7.1. Diversity-driven selection of exploration strategies in multi-armed bandits

In [40], we considered a scenario where an agent has multiple available strategies to explore an unknown environment. For each new interaction with the environment, the agent must select which exploration strategy to use. We provide a new strategy-agnostic method that treat the situation as a Multi-Armed Bandits problem where the reward signal is the diversity of effects that each strategy produces. We test the method empirically on a simulated planar robotic arm, and establish that the method is both able discriminate between strategies of dissimilar quality, even when the differences are tenuous, and that the resulting performance is competitive with the best fixed mixture of strategies.

7.2.7.2. Behavioral Diversity Generation in Autonomous Exploration Through Reuse of Past Experience

The production of behavioral diversity—producing a diversity of effects—is an essential strategy for robots exploring the world when facing situations where interaction possibilities are unknown or non-obvious. It allows to discover new aspects of the environment that cannot be inferred or deduced from available knowledge. However, creating behavioral diversity in situations where it is the most crucial, i.e. new and unknown ones, is far from trivial. In particular in large and redundant sensorimotor spaces, some effects can typically only be produced by a few number of specific motor commands. We introduced a method to create behavioral diversity by re-enacting past experiences, along with a measure that quantifies this diversity. We showed that our method is robust to morphological and representation changes, that it can learn how to interact with an object by reusing experience from another and how scaffolding behaviors can emerge by simply switching the attention of the robot to different parts of the environment. Finally, we showed that the method can robustly use simulated experiences and crude cognitive models to provide behavioural diversity in the real world. This result are under review.

7.3. Autonomous And Social Perceptual Learning

Participants: David Filliat [correspondant], Freek Stulp, Celine Craye, Yuxin Chen, Clement Masson, Adrien Matricon.

7.3.1. Incremental Learning of Object-Based Visual Saliency

Searching for objects in an indoor environment can be drastically improved if a task-specific visual saliency is available. We describe a method to learn such an object-based visual saliency in an intrinsically motivated way using an environment exploration mechanism. We first define saliency in a geometrical manner and use this definition to discover salient elements given an attentive but costly observation of the environment. These elements are used to train a fast classifier that predicts salient objects given large-scale visual features. In order to get a better and faster learning, we use intrinsic motivation to drive our observation selection, based on uncertainty and novelty detection. Our approach has been tested on RGB-D images, is real-time, and outperforms several state-of-the-art methods in the case of indoor object detection. We published these results in two conferences [43], [42]

7.3.2. Cross-situational noun and adjective learning in an interactive scenario

Learning word meanings during natural interaction with a human faces noise and ambiguity that can be solved by analysing regularities across different situations. We propose a model of this cross-situational learning capacity and apply it to learning nouns and adjectives from noisy and ambiguous speeches and continuous visual input. This model uses two different strategy: a statistical filtering to remove noise in the speech part and the Non Negative Matrix Factorization algorithm to discover word-meaning in the visual domain. We present experiments on learning object names and color names showing the performance of the model in real interactions with humans, dealing in particular with strong noise in the speech recognition. We published these results in a conference paper [41]

7.3.3. Learning representation with gated auto-encoders

We investigated algorithms that would be able to learn relevant visual or multi-modal features from data recorded while the robot performed some task. Representation learning is a currently very active research field, mainly focusing on deep-learning, which investigates how to compute more meaningful features from the raw high dimensional input data, providing a more abstract representation from which it should be easier to make decision or deduction (e.g classification, prediction, control, reinforcement learning). In the context of robotics, it is notably interesting to apply representation learning in a temporal and multi-modal approach exploiting vision and proprioception so as to be able to find feature that are relevant for building models of the robot itself and of its actions and their effect on the environment. Among the many existing approaches, we decided to explore the use of gated auto-encoders, a particular kind of neural networks including multiplicative connections, as they seem well adapted to this problem. Preliminary experimentations have been carried out with gated auto-encoders to learn transformations between two images. We observed that Gated Auto-Encoders (GAE) can successfully find compact representations of simple transformations such as translations, rotation or scaling between two small images. This is however not directly scalable to realistic images such as ones acquired by a robot's camera because of the number of parameters, memory size and computational power it would require (unless drastically downsampling the image which induces sensible loss of information). In addition, the transformation taking an image to the next one can be the combination of transformations due to the movement of several object in the field of view, composed with the global movement of the camera. This induces the existence of an exponential number of possible transformations to model, for which the basic GAE architecture is not suited. To tackle both issue, we are developing a convolutional architectures inspired from Convolutional Neural Networks (CNNs) that provide different modelisations for different parts of the image, which might be usefull to model combinations of transformations. Our Convolutional Gated Auto-Encoder is designed to perform generic feature learning in an unsupervised way (while most CNNs are trained in a supervised fasion) and we are currently testing it on realistic image sequences. We plan to extend this architecture to find relations between modalities as, for instance, proprioceptive information and its evolution could be used to predict the next visual features. Similarly, proprioceptive information could be used as a supervising signal to learn visual features.

7.3.4. Learning models by minimizing complexity

In machine learning, it is commonly assumed that simpler models have better chances at generalizing to new, unseen data. Following this principle, we developped an algorithm relying on minimization of a given complexity measure to build a collection of models which jointly explain the observation of the training datapoints. The resulting collection is composed of as few models as possible, each using as few dimensions as possible and each as regular as possible. As of now, each model is a multivariate polynomial, with the complexity of a polynomial of degree N in d variables being $N*d+1$. The complexity of the collection is the sum of the complexity of all its models. The algorithm starts by associating each datapoint to a local model of complexity 1 (degree 0, no variables), then models are iteratively merged into models of higher complexity, as long as those merges don't increase the complexity of the collection and as long as the resulting models stay within a certain distance of their associated datapoints. We applied this algorithm to the problem of inverse dynamics, which we studied in simulation. For a given robot, torques needed to compensate gravity at equilibrium are entirely determined by the values of its joint angles. As it is common that robots actually perform only low-dimensional tasks, and do not explore their full state space during normal operation, we would like the complexity of our models to mirror the structure of the task. When the task was expressed in the joint space, we got satisfying results on that point, and got good predictions for unseen datapoints. When the task was expressend in end-effector position, it turned out to be impossible to learn the underlying manifolds because a given end-effector position can correspond to various joint configurations, and thus to various torques, making it impossible to predict those torques from the end-effector position alone. We are currently working on applying this model to data generated by an exploration algorithm on a robot arm manipulating objects.

7.4. Applications for Robotic myoelectric prostheses: co-adaptation algorithms and design of a 3D printed robotic arm prosthesis

Participants: Pierre-Yves Oudeyer [correspondant], Manuel Lopes, Joel Ortiz, Mathilde Couraud, Aymar de Rugy, Daniel Cattaert, Florent Paclet.

Together with the Hybrid team at INCIA, CNRS, the Flowers team continued to work on establishing the foundations of a long-term project related to the design and study of myoelectric robotic prosthesis. The ultimate goal of this project is to enable an amputee to produce natural movements with a robotic prosthetic arm (open-source, cheap, easily reconfigurable, and that can learn the particularities/preferences of each user). This will be achieved by 1) using the natural mapping between neural (muscle) activity and limb movements in healthy users, 2) developing a low-cost, modular robotic prosthetic arm and 3) enabling the user and the prosthesis to co-adapt to each other, using machine learning and error signals from the brain, with incremental learning algorithms inspired from the field of developmental and human-robot interaction. In particular, in 2015 two lines of work were achieved, concerning two important scientific challenges, and in the context of a PEPS CNRS project:

First, an experimental setup was designed to allow fast prototyping of 3D printed robotic prostheses (internship of Joel Ortiz). This work was based on the use of the Poppy open-source modular platform, and resulted in a functional prototype. Several video demonstrations are available at: <https://forum.poppy-project.org/t/real-time-control-of-a-prosthetic-robotic-arm-with-muscle-activities/1656>.

Second, first versions of co-adaptation algorithms were designed, implemented and tested with human subjects, based on the combination of advanced models of the arm biomechanics and incremental learning algorithms (internship of Mathilde Couraud). An article is under preparation.

7.5. Applications for Educational Technologies

7.5.1. KidLearn

Participants: Manuel Lopes [correspondant], Pierre-Yves Oudeyer, Didier Roy, Benjamin Clement.

Kidlearn is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

Intelligent Tutoring System (ITS) are computer environments designed to guide students in their learning. Through the proposal of different activities, it provides teaching experience, guidance and feedback to improve learning. The FLOWERS team has developed several computational models of artificial curiosity and intrinsic motivation based on research on psychology that might have a great impact for ITS. Results showed that activities with intermediate levels of complexity, neither too easy nor too difficult but just a little more difficult than the current level, provide better teaching experiences. The system is based on the combination of three approaches. First, it leverages Flowers team's recent models of computational models of artificial curiosity and intrinsic motivation based on research in psychology and neuroscience. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. In 2014, we have run a second pilot experiment in elementary schools of Région Aquitaine, where 7-8 year old kids could learn elements of mathematics thanks to an educational software that presented the right exercises at the right time to maximize learning progress. [29]

7.5.2. Poppy System

Participants: Matthieu Lapeyre [correspondant], Nicolas Rabault, Pierre Rouanet, Pierre-Yves Oudeyer.

In the Poppy project we are working on the Poppy System which is a new modular and open-source robotic architecture. It is design to help people create and build custom robots. It permits, in a similar approach as Lego, building robots or smart objects using standardized elements.

Poppy System is an unified system where each essential robotic components (actuators, sensors, ...) is an independant module, connected with other through standardized interfaces.

- Unified mechanical interfaces which simplifies the assembly process and the design of 3D printable parts.
- Unified communication between elements using the same connector and bus for each module.
- Unified software makes it easy to program each module independantly.

The current Poppy robots (Humanoid, Torso, Ergo) will be updated using this novel architecture.

Our ambition is to create an ecosystem around this system so communities can develop custom modules, following the Poppy System standards, which can be compatible with all other Poppy robots.

7.5.3. Poppy Education

Participants: Pierre-Yves Oudeyer [correspondant], Didier Roy, Théo Segonds, Stéphanie Noirpoudre, Marie Demangeat, Thibault Desprez, Matthieu Lapeyre, Pierre Rouanet, Nicolas Rabault.

Poppy Education aims to create, evaluate and disseminate pedagogical kits “turnkey solutions” complete, open-source and low cost, for teaching computer science and robotics. It is designed to help young people to take ownership with concepts and technologies of the digital world, and provide the tools they need to allow them to become actors of this world, with a considerable socio-economic potential. It is carried out in collaboration with teachers and several official french structures (French National Education, Highschools, engineers schools, ...). For secondary education and higher education, scientific literacy centers, Fablabs.

The Poppy robotic platform used in the project is free hardware and software, printed in 3D, and is intended primarily for:

- learning of computer science and robotics,
- introduction to digital manufacturing (3D printing ...)
- initiation to the integration of IT in physical objects in humanoid robotics, mechatronics.
- artistic activities.

Educational sectors covered by the project are mainly: Enseignement d’exploration ICN en seconde, enseignement ISN en terminale S et bientôt en 1ère , filière STI2D, MPS seconde.

Users and their needs are placed at the center of this project. The pedagogical tools of the project are being created directly with them and evaluated in real life by experiments. Poppy Education is based on the robotic platform poppy, from which it is possible to construct different robots, including:

- Poppy Humanoid is a robust and complete robotics platform designed for genuine experiments in the real world and can be adapted to specific user needs.
- Poppy Torso is a variant of Poppy Humanoid. It is a torso humanoid robot that can be easily installed on a table.
- Poppy Ergo Jr is a robotic arm. Solid and inexpensive, it is perfect to be used in class. Poppy robots are easy to program. Different options are possible based on students level and teaching objectives :
- Pixl is a board who manage power and communication between a raspberry pi and robotis XL320 low cost motors. We use this bord for all our low cost robots.
- Python. Directly from a web browser, using Ipython notebooks (an interactive terminal, in a web interface for the Python Programming Language).
- Snap. The visual programming system Snap, which is a variant of Scratch. Its features allow a thorough introduction of IT.
- C ++, Java, Matlab, Ruby, Javascript, etc. thanks to a REST API that allows you to send commands and receive information from the robot with simple HTTP requests.

Poppy Humanoid, Torso and Ergo robots can be simulated with the free simulator V-REP. It is possible in the classroom to work on the simulated model and then allow students to run their program on the physical robot.

Experimentations have began to be setup in 10 high-schools of Region Aquitaine, and 3 university level institutions: Lycée Camille Jullian (Bordeaux), Lycée Victor Louis (Talence), Lycée Saint Genès (Talence), Lycée François Mauriac (Bordeaux), Lycée Jean Moulin (Langon), Lycée des Graves (Gradignan), Lycée Sud Medoc (Le Taillan Medoc), Lycée Alfred Kastler (Talence), Lycée Raoul Follereau (Nevers), Aérocampus Auqitaine, ENSEIRB/IPB, ENSAM Talence.

7.5.4. IniRobot : Education and Thymio II Project (partnership with EPFL)

Participants: Didier Roy [correspondant], Pierre-Yves Oudeyer.

IniRobot Project consists to produce and diffuse a pedagogical kit for teachers and animators, to help to train them directly or by the way of external structures. The aim of the kit is to initiate children to computer science and robotics. The kit provides a micro-world for learning, and takes an enquiry-based educational approach, where kids are led to construct their understanding through practicing an active investigation methodology within teams. It is based on the use of the Thymio II robotic platform. More details about this projects were published in RIE 2015 [50], which presents the detailed pedagogical objectives and a first measure of results showing that children acquired several robotics-related concepts. See also <https://dm1r.inria.fr/inirobot> or <http://www.inirobot.fr>. The project is carried out in main collaboration with the LSRO Laboratory from EPFL (Lausanne) and others collaborations with French National Education/Rectorat d' Aquitaine.

Deployment: After 16 months of activity, IniRobot is used by about 900 adults and 8000 children in 35 cities of France. Example of action in university: MEEF teacher training for the hope of Aquitaine. Example of action in school: training of all Gironde Pedagogical ICT Advisors, covering nearly 1000 schools. Example of action in the extracurricular time: training 82 facilitators TAP cities of Talence, Pessac, Lille, ..., CDC Gates of inter-seas. Example of national action: Training of the digital mediators of the 8 Inria centers.

GRAPHDECO Project-Team

7. New Results

7.1. Computer-Assisted Design with Heterogeneous Representations

7.1.1. *BendFields: Regularized Curvature Fields from Rough Concept Sketches*

Participants: Adrien Bousseau, Emmanuel Iarussi.

Designers frequently draw curvature lines to convey bending of smooth surfaces in concept sketches. We present a method to extrapolate curvature lines in a rough concept sketch, recovering the intended 3D curvature field and surface normal at each pixel of the sketch (Fig. 4). This 3D information allows us to enrich the sketch with 3D-looking shading and texturing. We first introduce the concept of *regularized curvature lines* that model the lines designers draw over curved surfaces, encompassing curvature lines and their extension as geodesics over flat or umbilical regions. We build on this concept to define the orthogonal cross field that assigns two regularized curvature lines to each point of a 3D surface. Our algorithm first estimates the projection of this cross field in the drawing, which is non-orthogonal due to foreshortening. We formulate this estimation as a scattered interpolation of the strokes drawn in the sketch, which makes our method robust to sketchy lines that are typical for design sketches. Our interpolation relies on a novel smoothness energy that we derive from our definition of regularized curvature lines. Optimizing this energy subject to the stroke constraints produces a dense non-orthogonal 2D cross field, which we then lift to 3D by imposing orthogonality. Thus, one central concept of our approach is the generalization of existing cross field algorithms to the non-orthogonal case. We demonstrate our algorithm on a variety of concept sketches with various levels of sketchiness. We also compare our approach with existing work that takes clean vector drawings as input.

This work is a collaboration with David Bommes from Titane project team at Inria Sophia-Antipolis, now at RWTH Aachen University. It has been published at ACM Transactions on Graphics (TOG) [7].

7.1.2. *Line Drawing Interpretation in a Multi-View Context*

Participant: Adrien Bousseau.

Many design tasks involve the creation of new objects in the context of an existing scene. Existing work in computer vision only provides partial support for such tasks. On the one hand, multi-view stereo algorithms allow the reconstruction of real-world scenes, while on the other hand algorithms for line-drawing interpretation do not take context into account. Our work combines the strength of these two domains to interpret line drawings of imaginary objects drawn over photographs of an existing scene (Fig. 5). The main challenge we face is to identify the existing 3D structure that correlates with the line drawing while also allowing the creation of new structure that is not present in the real world. We propose a labeling algorithm to tackle this problem, where some of the labels capture dominant orientations of the real scene while a free label allows the discovery of new orientations in the imaginary scene. We illustrate our algorithm by interpreting line drawings for urban planing, home remodeling, furniture design and cultural heritage.

This work is a collaboration with Jean-Dominique Favreau and Florent Lafarge from Titane project team, Inria Sophia-Antipolis. It has been published at the Conference on Computer Vision and Pattern Recognition (CVPR) [14].

7.1.3. *WrapIt: Computer-Assisted Crafting of Wire Wrapped Jewelry*

Participants: Adrien Bousseau, Emmanuel Iarussi.

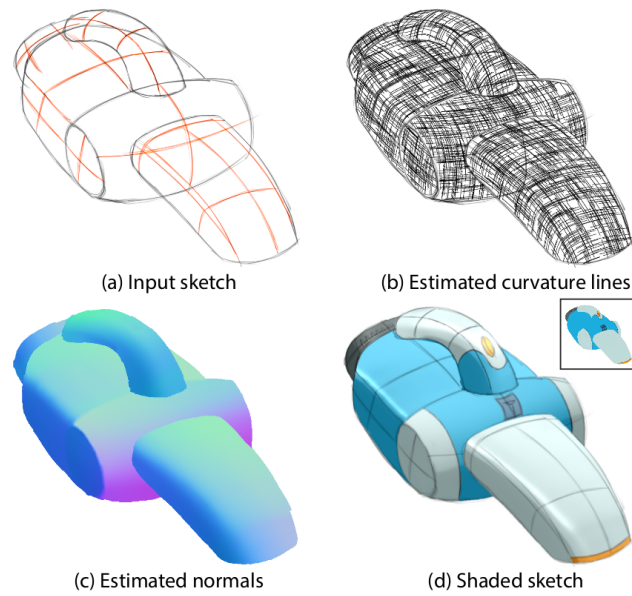


Figure 4. Our method [7] takes as input a rough design sketch with annotated curvature lines (a). We propose a novel smoothness energy to propagate the curvature information to all pixels (b), which allows us to recover surface normals (c) and compute shading (d).

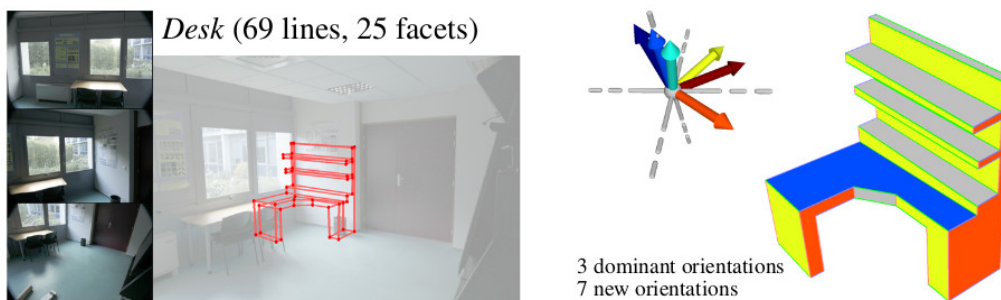


Figure 5. Our method [14] takes as input several photographs of a scene, along with a line drawing of a new object (left). We exploit the dominant orientations of the existing scene to reconstruct the line drawing in 3D (right).

Wire wrapping is a traditional form of handmade jewelry that involves bending metal wire to create intricate shapes. The technique appeals to novices and casual crafters because of its low cost, accessibility and unique aesthetic. We present a computational design tool that addresses the two main challenges of creating 2D wire-wrapped jewelry: decomposing an input drawing into a set of wires, and bending the wires to give them shape (Fig. 6). Our main contribution is an automatic wire decomposition algorithm that segments a drawing into a small number of wires based on aesthetic and fabrication principles. We formulate the task as a constrained graph labeling problem and present a stochastic optimization approach that produces good results for a variety of inputs. Given a decomposition, our system generates a 3D-printed custom support structure, or *jig*, that helps users bend the wire into the appropriate shape. We validated our wire decomposition algorithm against existing wire-wrapped designs, and used our end-to-end system to create new jewelry from clipart drawings. We also evaluated our approach with novice users, who were able to create various pieces of jewelry in less than half an hour.

This work is a collaboration with Wilmot Li from Adobe, San Francisco. The project was initiated by a 3-months visit of Emmanuel Iarussi at Adobe. It has been published at ACM Transactions on Graphics (Proc. SIGGRAPH Asia) [8].



Figure 6. Our system [8] helps novices convert a line drawing (left) into a real piece of jewelry (right).

7.1.4. How Novices Sketch and Prototype Hand-Fabricated Objects

Participant: Adrien Bousseau.

We are interested in how to create digital tools to support informal sketching and prototyping of objects by *novices*. Achieving this goal first requires a deeper understanding of how novices currently generate, explore, and communicate design ideas with traditional tools, i.e., sketches on paper and hands-on prototyping materials. We describe a study framed around two all-day design charrettes where participants perform a complete design process including ideation sketching, concept development and presentation, fabrication planning documentation and collaborative fabrication of hand-crafted prototypes. This structure allows us to control key aspects of the design process while collecting rich data about creative tasks, including sketches on paper, physical models, and videos of collaboration discussions. We observed that while participants had no formal training in design, they made use of advanced visualization techniques to convey 3D concepts. Participants also extensively used physical materials (paper, foam, cardboard) both to support concept exploration and to communicate their ideas to collaborators. We deduce from these observations recommendations for the conception of design tools adapted to the needs and skills of novices.

This work is a collaboration with Wendy McKay, Theophanis Tsandilas and Lora Oehlberg from the InSitu project team - Inria Saclay, in the context of the ANR DRAO project. It is conditionally accepted to ACM CHI 2016.

7.1.5. Vectorizing Rough Line Drawings

Participant: Adrien Bousseau.

Our goal in this project is to convert rough, freehand bitmap sketches to clean vector drawings, keeping three main objectives in mind: (i) the vectorial curves should approximate well the input drawing, (ii) the drawing should be composed of a small number of curves with few control points to preserve the compactness and editability of vector graphics, and (iii) the algorithm should support user guidance to disambiguate the multiple interpretations inherent to artistic inputs. Unfortunately, existing vectorization algorithms only partly satisfy these requirements. In particular, while most methods employ curve fitting to satisfy the first objective of data fidelity, this fitting is performed locally and is often sub-optimal with respect to our second objective of *low complexity*. To achieve our objectives, we propose to cast line drawing vectorization as a global optimization that balances data fidelity with model complexity. We express data fidelity as the goodness of fit of Bézier curve segments, and we express model complexity as the number and degree of curve segments that compose the output drawing. Our algorithm produces clean, compact and editable vector drawings from bitmap sketches.

This ongoing work is a collaboration with Jean-Dominique Favreau and Florent Lafarge from Titane project team, Inria Sophia-Antipolis.

7.1.6. Exploring Design Spaces with Sketch-Based Rendering

Participant: Adrien Bousseau.

Designers often start product design by drawing many quick and imperfect sketches. These sketches typically capture shape variations of a concept from different viewpoints. We introduce *sketch-based rendering* as a way to help designers explore the design space induced by such sketches. Our interactive tool allows designers to interpolate between the sketches, providing a continuous, 3D-like visualization of the concept and its variations without requiring explicit 3D information.

We propose an iterative algorithm to match and warp between sketches using little user interaction. We designed this algorithm to address the specific challenges inherent to concept sketches, in particular the fact that they are dominated by contours rather than color or texture, and that these contours should not distort during interpolation. We also describe how to approximate the relative camera positions of different sketches from the magnitude of their 2D motion fields. This approach allows plausible 3D-like camera movements despite the presence of sketch distortions and variations that prevent standard camera calibration. Our tool, thus, fills a gap in the initial stage of the product design pipeline by allowing designers and their patrons to make better informed choices before proceeding to more expensive 3D modeling and prototyping.

This ongoing work is a collaboration with Ishan Darolia and Vinay Namboodiri from IIT Kampur and Rahul Arora and Karan Singh from University of Toronto.

7.1.7. Sketch-Based Inverse Procedural Modeling

Participant: Adrien Bousseau.

Designing and modeling 3D objects is a crucial skill in various areas of entertainment, science, and engineering. However, this task is notoriously hard and unintuitive, especially for novices. Prior work has addressed the modeling problem from many different directions. Sketch-based modeling exploits human intuition and experience in drawing objects. Nevertheless, the quality of the final 3D model depends on the sketching skills of the user, the amount of details added to the drawing, and ability to resolve inherent ambiguities of the sketching process. Another popular direction is procedural modeling, which has been successfully used to create detailed and complex cities, realistic and growing vegetation, and other man-made objects. But procedural modeling is difficult to control and thus hard to use as an exploratory design tool making it accessible only to experts. Our goal in this project is to leverage both the intuitiveness, freedom and flexibility of sketching and the precision, exactness, and detail amplification of procedural modeling. Users of our system begin to sketch a 3D model using a mouse or a digital pen on a tablet. After only a few strokes, our algorithm finds a compact 3D procedural representation that matches the sketch while augmented it with geometric details.

This ongoing work is a collaboration with Gen Nishida, Bedrich Benes, Ignacio Garcia Dorado and Daniel Aliaga from Purdue University.

7.1.8. A data-based approach to retrieve the viewpoint of a design sketch

Participants: Johanna Delanoy, Adrien Bousseau.

Designing objects requires frequent transitions from a 2D representation, the sketch, to a 3D one. Because 3D modeling is time consuming, it is made only during late phases of the design process. Our long term goal is to allow designers to automatically generate 3D models from their sketches. In this work, we address the preliminary step of recovering the viewpoint under which the object is drawn. We adopt a data-driven approach where we build correspondences between the sketch and 3D objects of the same class from a database. In particular, we relate the curvature lines and contours of the 3D objects to similar lines commonly drawn by designers using histograms of orientation. The 3D objects from the database are then used to vote for the viewpoints and the more probable ones are chosen. Our results on design sketches suggest that using both contours and curvature lines give higher precision than using either one. In particular, curvature information improves viewpoint retrieval when details of the objects are different from the sketch.

The work has been published in the journal *Revue Française d'Informatique Graphique* and presented at the 28th Journées de l'Association Française d'Informatique Graphique [5].

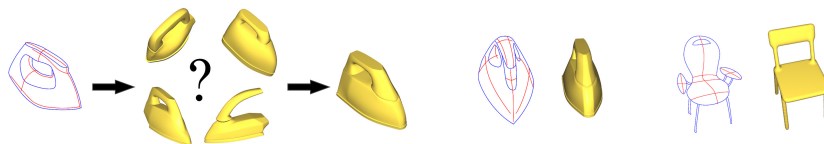


Figure 7. Our method [5] allows to retrieve the viewpoint of a design sketch, using a collection of 3D objects.

7.2. Graphics with Uncertainty and Heterogeneous Content

7.2.1. Multi-View Intrinsic Images for Outdoors Scenes with an Application to Relighting

Participants: Sylvain Duchêne, Clement Riant, Gaurav Chaurasia, Stefan Popov, Adrien Bousseau, George Drettakis.

We introduce a method to compute intrinsic images for a multi-view set of outdoor photos with cast shadows, taken under the same lighting (Fig. 8). We use an automatic 3D reconstruction from these photos and the sun direction as input and decompose each image into reflectance and shading layers, despite the inaccuracies and missing data of the 3D model. Our approach is based on two key ideas. First, we progressively improve the accuracy of the parameters of our image formation model by performing iterative estimation and combining 3D lighting simulation with 2D image optimization methods. Second we use the image formation model to express reflectance as a function of discrete visibility values for shadow and light, which allows us to introduce a robust visibility classifier for pairs of points in a scene. This classifier is used for shadow labelling, allowing us to compute high quality reflectance and shading layers. Our multi-view intrinsic decomposition is of sufficient quality to allow relighting of the input images. We create shadow-caster geometry which preserves shadow silhouettes and using the intrinsic layers, we can perform multi-view relighting with moving cast shadows. We present results on several multi-view datasets, and show how it is now possible to perform image-based rendering with changing illumination conditions.

This work was published in *ACM Transactions on Graphics* [2].



Figure 8. Our algorithm [2] decomposes outdoor images (a) into reflectance (c) and shading (d). This decomposition enables relighting of photographs by rendering new shadows in the shading layer (b).

This work is part of an industrial partnership with Autodesk and has been published in ACM Transactions on Graphics [2].

7.2.2. A Bayesian Approach for Selective Image-Based Rendering using Superpixels

Participants: Rodrigo Ortiz Cayon, Abdelaziz Djelouah, George Drettakis.

Many recent Image-Based Rendering (IBR) algorithms have been proposed each having different strengths and weaknesses, depending on 3D reconstruction quality and scene content. Each algorithm operates with a set of hypotheses about the scene and the novel views, resulting in different quality/speed trade-offs in different image regions. We developed a principled approach to select the algorithm with the best quality/speed trade-off in each region. To do this, we propose a Bayesian approach, modeling the rendering quality, the rendering process and the validity of the assumptions of each algorithm. We then choose the algorithm to use with Maximum a Posteriori estimation. We demonstrate the utility of our approach on recent IBR algorithms which use oversegmentation and are based on planar reprojection and shape-preserving warps respectively. Our algorithm selects the best rendering algorithm for each superpixel in a preprocessing step; at runtime our selective IBR uses this choice to achieve significant speedup at equivalent or better quality compared to previous algorithms. The work has been published in the International Conference on 3D Vision (3DV) - 2015 [15].

7.2.3. Uncertainty Modeling for Principled Interactive Image-Based Rendering

Participants: Rodrigo Ortiz Cayon, George Drettakis.

Despite recent advances in IBR methods, they are limited in regions of the scene which are badly or completely unreconstructed. Such regions have varying degrees of uncertainty, which previous solutions treat with heuristic methods. Currently we attempt to develop a comprehensive model of uncertainty for interactive IBR. Regions with high uncertainty would feed an iterative multi-view depth synthesis algorithm. For the rendering we will formalize an unified IBR algorithm, which provides a good quality/speed tradeoff by combining the



Figure 9. In top-left, we use planes fronto-parallel to the input view which fail for trees and slanted planes. Using local plane estimation top-right the result is improved, especially for slanted planes (blue box). Using the shape preserving warp bottom-left of the warping method we previously developed, better results are achieved for the tree (red box), but the quality of the slanted planes is worse. Our algorithm [15] bottom-right makes the correct choice locally, giving the best solution in each case.

advantages of forward warping and depth-based backprojection algorithms and includes plausible stereoscopic rendering for unreconstructed volumetric regions.

7.2.4. *Multi-view Inpainting*

Participants: Theo Thonat, George Drettakis.

We are developing a new approach for removing objects in multi-view image datasets. For a given target image from which we remove objects, we use Image-Based Rendering for reprojecting the other images into the target and for regions not visible in any other image we use inpainting techniques. The difficulties reside in formalizing the unified algorithm and enforcing multi-view consistency. This is an ongoing project in collaboration with Adobe Research (E. Shechtman and S. Paris).

7.2.5. *Beyond Gaussian Noise-Based Texture Synthesis*

Participants: Kenneth Vanhoey, Georgios Kopanas, George Drettakis.

Texture synthesis methods based on noise functions have many nice properties: they are continuous (thus resolution-independent), infinite (can be evaluated at any point) and compact (only functional parameters need to be stored). A good method is also non-repetitive and aperiodic. Current techniques, like Gabor Noise, fail to produce structured content. They are limited to so-called “Gaussian textures”, characterized by second-order statistics like mean and variance only. This is suitable for noise-like patterns (e.g., marble, wood veins, sand) but not for structured ones (e.g., brick wall, mountain rocks, woven yarn). Other techniques, like Local Random-Phase noise, leverage some structure but as a trade-off with repetitiveness and periodicity.

In this project, we model higher-order statistics produced by noise functions. Then we define an algorithm for maximal-entropy sampling of the parameters of the noise functions so as to meet prescribed statistics to reproduce. This sampling ensures both the reproduction of higher-order visual features with high probability, like edges and ridges, and non-repetitiveness plus aperiodicity thanks to the stochastic sampling method. We are currently investigating a learning method so as to inject into the model the appropriate prescribed statistics deduced from an input exemplar image.

This ongoing work is a collaboration with Ian Jermyn from Durham University and will be submitted for publication in 2016.

7.2.6. *Unifying Color and Texture Transfer for Predictive Appearance Manipulation*

Participants: Fumio Okura, Kenneth Vanhoey, Adrien Bousseau, George Drettakis.

Recent color transfer methods use local information to learn the transformation from a Source to an Exemplar image, and then transfer this appearance change to a Target image (figure 10 (a) to (d)). These solutions achieve successful results for general mood changes, e.g., changing the appearance of an image from “sunny” to “overcast”. However, they fail to create new image content, such as leaves on a bare tree (figure 10 (d)). Texture transfer, on the other hand, can synthesize such content but tends to destroy image structure (figure 10 (e)). We propose the first algorithm that unifies color and texture transfer, outperforming both by automatically leveraging their respective strengths (figure 10 (f)). A key novelty in our approach resides in teasing apart appearance changes that can be modeled simply as changes in color versus those that require new image content to be generated. Our method starts with an analysis phase which evaluates the success of color transfer on the Source/Exemplar scene. To do so, color transfer parameters are learned on this pair, and applied on the Source. The color transferred Source image is then evaluated against the Exemplar which serves as a ground truth, using texture distance metrics (textons in our case). This provides information on the localization of success and failure of color transfer on this scene. This analysis then drives the synthesis: a selective, iterative texture transfer algorithm that simultaneously predicts the success of color transfer on the Target and synthesizes new content using texture transfer where needed. Synthesis exploits a dense pixel matching between the Source/Exemplar scene, on which information is learned, and the Target/Output scene, on which we want to synthesize. The algorithm iterates between synthesizing the new scene by locally using either color or texture transfer, and improving the dense matching on the scene being synthesized. As a result, it leverages the best of both techniques on a variety of scenes by transferring large temporal changes between photographs, such as change of season and flooding. We demonstrate this with seasonal changes on vegetation (e.g., trees) and snow, and on examples involving flooding.

This work is a collaboration with Alexei Efros from UC Berkeley in the context of the associate team CRISP. It has been published in Computer Graphics Forum [9] and was accepted and presented at the Eurographics symposium on Rendering.

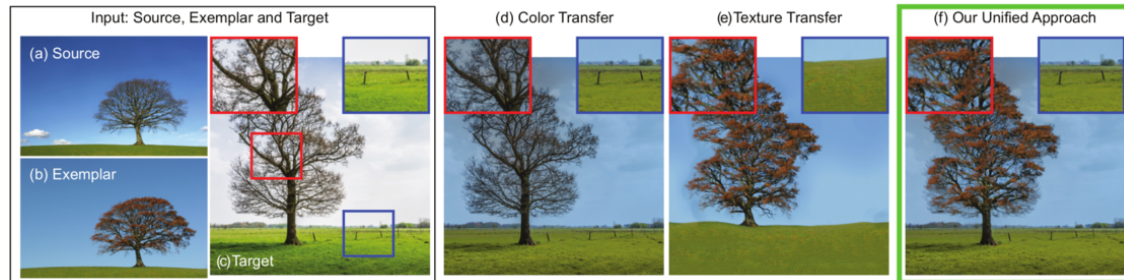


Figure 10. Illustration of our appearance prediction method [9]. The future appearance of a target image (c) is predicted (f) based on the knowledge learned from a quasi-aligned source-exemplar pair ((a) and (b)) which characterizes an analogous transformation. The key insight is to selectively operate color transfer ((d): only operate rigid local color histogram transformations, i.e., change the background's overall mood) or texture transfer ((e): copying pixels or patches from the exemplar, i.e. synthesize the tree's leaves) where suitable, so as to obtain an improved result (f).

7.2.7. Simplification of Triangle Meshes with Digitized Radiance

Participant: Kenneth Vanhoey.

Very accurate view-dependent surface color of virtual objects can be represented by outgoing radiance of the surface. This data forms a surface light field, which is inherently 4-dimensional, as the color is varying both spatially and directionally. Acquiring this data and reconstructing a surface light field of a real-world object can result in very large datasets, which are very realistic, but tedious to store and render. In this project, we consider the processing of outgoing radiance stored as a vertex attribute of triangle meshes, and especially propose a principled simplification technique. We show that when reducing the global memory footprint of such acquired objects, smartly reducing the spatial resolution, as opposed to the directional resolution, is an effective strategy for overall appearance preservation. To define such simplification, we define a new metric to guide an iterative edge collapse algorithm. Its purpose is to measure the visual damage introduced when operating a local simplification. Therefore, we first derive mathematical tools to calculate with radiance functions on the surface: interpolation, gradient computation and distance measurements. Then we derive a metric using these tools. We particularly ensure that the mathematical interpolation used in the metric is coherent with the non-linear interpolation we use for rendering, which makes the math coherent with the rendered object. As a result we show that both synthetic and acquired objects benefit from our radiance-aware simplification process: at equal memory footprint, visual quality is improved compared to state of the art alternatives.

This work is a collaboration with the ICube laboratory, Strasbourg, France. It was published in the Computer Graphics Journal [11] and was accepted and presented at Computer Graphics International 2015 in Strasbourg, France.

7.2.8. Video based rendering for vehicles

Participants: Abdelaziz Djelouah, Georgios Koulieris, George Drettakis.

The main objective of image based rendering methods is to provide high quality free-view point navigation in 3D scenes using only a limited set of pictures. Despite the good visual quality achieved by most recent methods, the results still look unrealistic because of the static nature of the rendered scenes. This project is in the general context of enriching image based rendering experience by adding dynamic elements and we are particularly interested by adding vehicles.

Vehicles represent an important proportion of the dynamic elements in any urban scene and adding an object with such a complex appearance model has many challenges. First, contrary to classic IBR the number of viewpoints is limited because all input videos must be recorded at the same time. Also, because of this limited number of viewpoints, using classic multi-view reconstruction methods does not produce good results. Instead we use 3D stock models as proxy for the cars. The first step is the registration of the 3D model with the input videos. Then, using the 3D model, the input videos are processed to extract the different visual layers (base color, reflections, transparency, etc.). Finally, the objective is to find the appropriate way to combine the 3D model and the extracted layers to provide the most realistic image from any viewpoint.

This ongoing work is a collaboration with Gabriel Brostow from University College London in the context of the CR-PLAY EU project and with Alexei Efros from UC Berkeley.

7.2.9. Finger-Based Manipulation in Immersive Spaces and the Real World

Immersive environments that approximate natural interaction with physical 3D objects are designed to increase the user's sense of presence and improve performance by allowing users to transfer existing skills and expertise from real to virtual environments. However, limitations of current Virtual Reality technologies, e.g., low-fidelity real-time physics simulations and tracking problems, make it difficult to ascertain the full potential of finger-based 3D manipulation techniques. This project decomposes 3D object manipulation into the component movements, taking into account both physical constraints and mechanics. We fabricate five physical devices that simulate these movements in a measurable way under experimental conditions. We then implement the devices in an immersive environment and conduct an experiment to evaluate direct finger-based against ray-based object manipulation. The key contribution of this work is the careful design and creation of physical and virtual devices to study physics-based 3D object manipulation in a rigorous manner in both real and virtual setups.

This work was presented at IEEE Symposium on 3D User Interfaces [12], and is in collaboration with the EXSITU Inria group in Paris (T. Tsandilas, W. Mackay, L. Oehlberg).



Figure 11. A user in our immersive environment (left) for finger-based manipulation [12]. Completing a 6 DoF manipulation task in real (center) and virtual (right) settings.

7.2.10. Gaze Prediction using Machine Learning for Dynamic Stereo Manipulation

Participants: Georgios Koulieris, George Drettakis.

Comfortable, high-quality 3D stereo viewing is becoming a requirement for interactive applications today. The main challenge of this project is to develop a gaze predictor in the demanding context of real-time, heavily task-oriented applications such as games. Our key observation is that player actions are highly correlated with the present state of a game, encoded by game variables. Based on this, we train a classifier to learn these correlations using an eye-tracker which provides the ground-truth object being looked at. The classifier is used at runtime to predict object category – and thus gaze – during game play, based on the current state of game variables. We use this prediction to propose a dynamic disparity manipulation method, which provides rich and comfortable depth. We evaluate the quality of our gaze predictor numerically and experimentally, showing that it predicts gaze more accurately than previous approaches. A subjective rating study demonstrates that our localized disparity manipulation is preferred over previous methods.

This is a collaboration with the Technical University of Crete (K. Mania) and Cottbus University (D. Cunningham), and will be presented at IEEE VR 2016.

7.2.11. Compiling High Performance Recursive Filters

Infinite impulse response (IIR) or recursive filters, are essential for image processing because they turn expensive large-footprint convolutions into operations that have a constant cost per pixel regardless of kernel size. However, their recursive nature constrains the order in which pixels can be computed, severely limiting both parallelism within a filter and memory locality across multiple filters. Prior research has developed algorithms that can compute IIR filters with image tiles. Using a divide-and-recombine strategy inspired by parallel prefix sum, they expose greater parallelism and exploit producer-consumer locality in pipelines of IIR filters over multi-dimensional images. While the principles are simple, it is hard, given a recursive filter, to derive a corresponding tile-parallel algorithm, and even harder to implement and debug it. We show that parallel and locality-aware implementations of IIR filter pipelines can be obtained through program transformations, which we mechanize through a domain-specific compiler. We show that the composition of a small set of transformations suffices to cover the space of possible strategies. We also demonstrate that the tiled implementations can be automatically scheduled in hardware-specific manners using a small set of generic heuristics. The programmer specifies the basic recursive filters, and the choice of transformation requires only a few lines of code. Our compiler then generates high-performance implementations that are an order of magnitude faster than standard GPU implementations, and outperform hand tuned tiled implementations of specialized algorithms which require orders of magnitude more programming effort – a few lines of code instead of a few thousand lines per pipeline. This work was presented at the High Performance Computing conference and is a collaboration with F. Durand, J. Ragan-Kelley and G. Chaurasia of MIT and S. Paris of Adobe [13].

7.2.12. Probabilistic Connections for Bidirectional Path Tracing

Participants: Sefan Popov, George Drettakis.

Bidirectional path tracing (BDPT) with Multiple Importance Sampling is one of the most versatile unbiased rendering algorithms today. BDPT repeatedly generates sub-paths from the eye and the lights, which are connected for each pixel and then discarded. Unfortunately, many such bidirectional connections turn out to have low contribution to the solution. The key observation in this project is that we can importance sample connections to an eye sub-path by considering multiple light sub-paths at once and creating connections probabilistically. We do this by storing light paths, and estimating probability mass functions of the discrete set of possible connections to all light paths. This has two key advantages: we efficiently create connections with low variance by Monte Carlo sampling, and we reuse light paths across different eye paths. We also introduce a caching scheme by deriving an approximation to sub-path contribution which avoids high-dimensional path distance computations. Our approach builds on caching methods developed in the different context of VPLs. Our Probabilistic Connections for Bidirectional Path Tracing approach raises a major challenge, since reuse results in high variance due to correlation between paths. We analyze the problem of path correlation and derive a conservative upper bound of the variance, with computationally tractable sample weights. We present results of our method which shows significant improvement over previous unbiased global illumination methods, and evaluate our algorithmic choices.

This work was in collaboration with R. Ramamoorthi (UCSD) and F. Durand (MIT) and appeared in the Eurographics Symposium on Rendering [10].

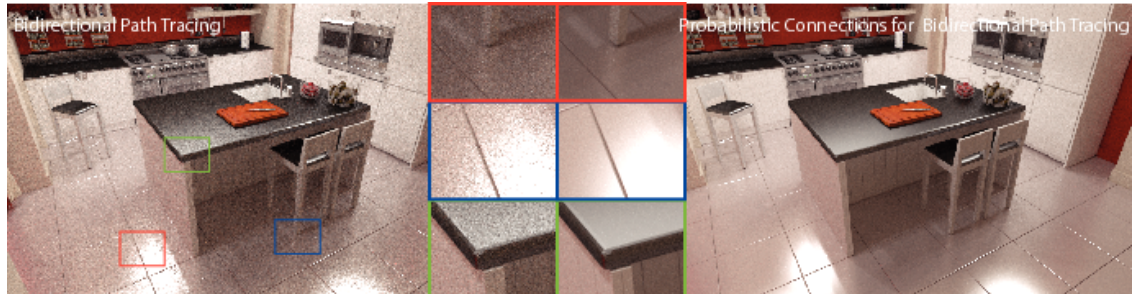


Figure 12. Our Probabilistic Connections for Bidirectional Path Tracing [10] approach importance samples connections to an eye sub-path, and greatly reduces variance, by considering and reusing multiple light sub-paths at once. Our approach (right) achieves much higher quality than bidirectional path-tracing on the left for the same computation time.

GRAPHIK Project-Team

7. New Results

7.1. Ontology-Based Query Answering with Existential Rules

Participants: Jean-François Baget, Meghyn Bienvenu, Fabien Garreau, Michel Leclère, Marie-Laure Mugnier, Swan Rocher, Federico Ulliana.

Since Meghyn Bienvenu joined the team very recently (September 2015), we only include here the work she did in collaboration with GraphIK members.

Ontology-based query answering (and more generally *Ontology-Based Data Access, OBDA*) is a new paradigm in data management, which takes into account inferences enabled by an ontology when querying data. In other words, the notion of a database is replaced by that of a knowledge base, composed of data (also called facts) and of an ontology. In this context, *existential rules* (also called Datalog+) have been proposed to represent the ontological component [42], [41]. This expressive formalism generalizes both description logics used in OBDA (such as \mathcal{EL} and DL-Lite) and Datalog, the language of deductive databases. Since about five years, we have been studying the theoretical foundations of this framework (mainly concerning decidability and complexity) and developing associated algorithmic techniques. We have started the development of a platform dedicated to OBDA with existential rules (see section 6.3).

Before presenting this year's results, we recall the two classical ways of processing rules, namely forward chaining and backward chaining, also known as "materialization" and "query rewriting" in the OBDA setting. In forward chaining, the rules are applied to enrich the initial data and query answering can then be solved by evaluating the query against the "saturate" database (as in a classical database system *i.e.*, with forgetting the rules). The backward chaining process can be divided into two steps: first, the initial query is rewritten using the rules into a first-order query (typically a union of conjunctive queries, UCQ); then the rewritten query is evaluated against the initial database (again, as in a classical database system). Since entailment is not decidable with general existential rules, both forward and backwards processes may not halt.

7.1.1. Embedding transitivity rules.

In recent years, many classes of existential rules have been exhibited for which CQ entailment is decidable. However, most of these classes cannot express transitivity of binary relations, a frequently used modelling construct. We began to investigate the issue of whether transitivity can be safely combined with decidable classes of existential rules. On the one hand, we obtained negative results, proving that transitivity is incompatible with many classes having finite chase, and with UCQ-reducible classes in general. Second, we showed that transitivity can be safely added to linear rules (a subclass of guarded rules, which generalizes the description logic DL-Lite_R) in the case of atomic CQs, and also for general CQs if we place a minor syntactic restriction on the rule set (only needed when predicate arity is strictly greater than 2). Finally, we pinpointed the combined and data complexities of query entailment over linear rules + transitivity.

▷ *IJCAI 2015* [22]

7.1.2. A generic algorithm for query reformulation.

We first designed and implemented a query reformulation algorithm that takes as input any set of existential rules and a UCQ q , and outputs a sound, minimal and complete UCQ-reformulation of q , whenever such a reformulation exists (*i.e.*, when the set of existential rules is UCQ-reducible). The core operation, unification, relies on a special technique that we first developed for conceptual graphs ("piece-unification"). A noteworthy feature of the implemented unification is that it is able to process rules without decomposing their head into single atoms. Experiments showed that this feature has a very high impact on the efficiency of query reformulation in terms of running time.

This algorithm can be seen as an instantiation of a generic reformulation algorithm, parametrized by a reformulation operator. As a complementary work, we studied the properties that should be fulfilled by any reformulation operator to ensure the correctness and the termination of this generic algorithm and analyzed some known operators with respect to these properties.

▷ *Semantic Web Journal 2015* [15]

7.1.3. Optimization of query reformulation algorithms

Query reformulation techniques have the advantage of being independent from the data. However, a main bottleneck is that the size of the obtained query can be exponential in the size of the original query, hence the produced reformulation maybe not usable in practice (and the corresponding SQL query may not even be accepted by the RDBMS). To overcome this combinatorial explosion in practice, we made two proposals, which have in common to consider other forms of reformulation, while staying equivalent to UCQs in terms of expressivity.

We defined semi-conjunctive queries (SCQs), which are a syntactical extension of conjunctive queries allowing for internal disjunctions. Briefly, a union of SCQs can be encoded in a more compact form than a UCQ. We designed and implemented an algorithm called Compact, which computes a sound and complete reformulation of a UCQ in the form of a union of SCQs (USCQ). First experiments showed that USCQs are both very efficiently computable and (often) more efficiently evaluable than their equivalent UCQs.

We developed another solution, which starts from a simple observation: in practice, combinatorial explosion is mainly due to some very simple rules, which form the core of any ontology. These rules typically express concept and relation hierarchies, concept properties and relation signatures. We proposed a technique that consists in compiling these rules into a preorder on atoms and embedding this preorder into the reformulation process. This allows us to compute compact reformulations that can be considered as “pivotal” representations, in the sense that they can be easily translated into different kinds of queries that can be evaluated by different kinds of database systems (e.g., unfolded into a classical UCQ or a USCQ, processed as such on data saturated by the compilable rules, or transformed into a Datalog program). Experiments show that this technique leads to substantial gains in the query reformulation process, in terms of size and runtime, it scales on very large ontologies (several ten thousands of rules), and it is competitive w.r.t. other existing tools, including those tailored for more specific rules corresponding to DL-Lite ontologies. This technique has been implemented in the software platform Graal.

▷ *IJCAI 2015*[28], *RuleML 2015*[23]

7.1.4. Ontology-based query answering with Semantic Web languages

On the one hand, we proposed Deductive RDF Triplestores, which are RDF knowledge bases equipped with Datalog inference rules. This work was developed in the context of the tool MyCorporisFabrica <http://www.mycorporisfabrica.org/>, an ontology-based tool for querying complex anatomical models.

In particular, we studied how to extract modules from deductive RDF triplestores. Indeed, many ontologies are extremely large, while users often need to reuse only a small part of resources in their work. A module is a Deductive RDF Triplestore entailed from the reference knowledge base, which is defined upon a restricted vocabulary. We proposed a new semantics for bounded-level modules allowing one to control their size, and then presented extraction algorithms compliant with the novel semantics.

▷ *AAAI 2015*[30] and *Journal of Biomedical Semantics* [16]. In collaboration with Marie-Christine Rousset (U. of Grenoble) and MyCorporisFabrica’s team.

On the other hand, in the context of the Graal platform, we defined a translation from the Semantic Web Ontological Language OWL 2 to our existential rule format. This gave rise to the definition of the “existential rule” OWL 2 profile, which covers the so-called tractable profiles of OWL 2 (see Section 6.3).

▷ *RuleML challenge*[33] (this paper obtained the RuleML 2015 challenge award)

7.2. Reasoning with Imperfect Information and Priorities

Participants: Abdallah Arioua, Patrice Buche, Madalina Croitoru, Jérôme Fortin, Nouredine Tamani, Rallou Thomopoulos.

This year, we mainly explored the use of argumentation frameworks in practical applications. Indeed, we have been involved in three main projects that have employed argumentation techniques. The projects were all in the context of agronomy where the nature of the problem studies fits well the use of argumentation: (1) the knowledge bases considered to model the domain are inconsistent, (2) the reasoning / decision making has to take into account the inconsistency, (3) the end user is a non-computing expert thus explanation facilities are needed.

We enumerate below the three projects and explain our results:

- Bread Making project (financed by the Food and Bioproducts department at INRA) investigates the possibility of using wholemeal flour in bread as opposed to classic white flour. The main theoretical result that we exploited here was the instantiation work into the existential rule framework done with Srdjan Vesic [43]. We used reverse engineering and the subsequent logic-based argumentation in order to provide the experts with a cartography of possible pros and cons of using one type of flour vs the other.
 - ▷ *Ecological Informatics 2015* [18]
- EcoBioCap (FP7 EU project led by INRA Montpellier, see Section 9.1) investigates the conception of biodegradable packaging for fruits and vegetables. The main theoretical result used here concerns the fuzzy aspects of argumentation but the modeling of the problem using argumentation and subsequent argument elicitation was also a very challenging process.
 - ▷ *Computers and Electronics in Agriculture 2015* (2 papers) [14] [17]
- DURDUR (ANR project led by INRA Montpellier, see Section 9.1) investigates the technological itineraries to grow durum wheat for subsequent pasta and couscous making. This ongoing project investigates the use of argumentation for explanation facilities.
 - ▷ Initial results have been published in *SUM 2015*[19] and *DEXA 2015* [20].

7.3. Quality and interoperability of large document catalogues

Participants: Michel Chein, Madalina Croitoru, Alain Gutierrez, Michel Leclère, Rallou Thomopoulos.

The work in this research line takes place in the ANR project Qualinca, devoted to methods and tools to repair linkage errors in bibliographical databases (see Qualinca in Section 9.1). Within this project, we specially work with our applicative partner ABES (French Agency for Academic Libraries, <http://www.abes.fr/>).

ABES manages several catalogues and authority bases, in particular the *Sudoc*, the collective catalogue of French academic libraries. ABES also provides services to libraries and end-users, as well as to other catalogue managers (e.g., OCLC for Worldcat and, in France, Adonis for the Isidore platform).

This year, we devoted most of our research effort to the following aspects in collaboration with ABES:

1. the finalization of a conceptual model of ABES librarian expertise in their linkage activity, and its formalization in our theoretical framework; the formalized model is both logical (the knowledge is expressed by facts, rules and constraints in first-order logic) and numerical (some predicates, which correspond to qualitative criteria, are computed by numerical functions, which themselves take as input the result of logical queries to the knowledge base).
2. the development of a diagnosis prototype, called SudoQual, which implements this model; in brief, SudoQual takes as input a given appellation (i.e., family name and first name), retrieves all references potentially associated with this appellation and outputs *sameAs* and *Different* links between these references. To develop SudoQual, we built an API on top of our tool Cogui.
3. first experiments with SudoQual on the Sudoc base, with the results being checked manually by ABES librarians.

▷ *Research report* [37]

This work required a tight collaboration with ABES (materialized by bimonthly meetings and numerous punctual exchanges). The first experiments yield extremely satisfactory results, hence ABES is now considering turning SudoQual into a production tool used by librarians in their daily work to validate/correct authority links in the Sudoc catalogue. This requires to define a suitable user-interface, which is an issue we are currently discussing with ABES. We are also preparing experiments at a larger scale on a sample provided by ABES.

Besides, in collaboration with Qualinca partner LRI, we developed a method and a tool to fusion data linked by “same-as” links. More precisely, given an RDF dataset, our tool allows to merge “same as” data, which are often conflictual, into a unified and consistant representation using a multi-criteria decision method. The tool was evaluated on a dataset provided by INA and LIG, two other partners of Qualinca.

▷ *EGC 2015* [32]

Still with the LRI partner, who developed a logic-based decision tool that statuates on the validity of same-as links in RDF data, we investigated the use of argumentation techniques to explain why “same-as” links are invalidated by this tool.

▷ *SUM 2015* [25]

HEPHAISTOS Project-Team

7. New Results

7.1. Robotics

7.1.1. Cable-driven parallel robots (CDPR)

7.1.1.1. Analysis of Cable-driven parallel robots

Participants: Alessandro Berti, Laurent Blanchet, Houssein Lamine, Jean-Pierre Merlet [correspondant], Yves Papegay, Rémy Ramadour.

We have continued the analysis of suspended CDPRs for control and design purposes. For control it is essential to determine the current pose of the robot for given cable lengths (forward kinematics, FK) and to be able to calculate the cable lengths for a given pose of the platform (inverse kinematics, IK). If the cables are supposed to be non-deformable the IK problem is trivial and has a single solution but the FK is complex, admits several solutions and raises several issues. We have shown in the past that to get all FK solutions for a CDPR with m cables we have to consider not only the case where all cables are under tension but also have to solve the FK for all combinations of cables under tension with 1 to m cables. Surprisingly the FK is more difficult if the CDPR has less than 6 cables under tension. Our team, in collaboration with M. Carricato of Bologna University, is the first to have designed a solving algorithm that allow to compute in a guaranteed manner all FK solutions [21], [22]. The FK problem is different if it is intended to be used in a real-time context as in that case we have the extra information of the platform pose a short time before. After a small change in the cable lengths we may assume a small change in the pose platform but using Newton method with the previous pose cannot guarantee to provide the current pose. We have proposed an algorithm that is guaranteed to get the current pose and is also able to determine if the CDPR may be sufficiently close to a singularity so that multiple solutions are possible [11]. However the assumption of a small change in the platform pose may not always hold, a point that we have shown theoretically and experimentally. We have then proposed an algorithm that uses a model of the coiling process to determine if a drastic change in the pose may occur between two sampling time [11] and also allows one to better estimate the cable tensions on a trajectory. We have for example shown that sudden and important changes in these tensions may occur. Another issue arises for non-deformable cables and CDPR with more than 6 cables in a suspend configuration. In the past we have shown that there always will be at most 6 cables under tension whatever the number of cables. For a given pose there may be several possible set of cables under tension (called *cable configuration*), each of them having different characteristics in terms of maximal tension, sensitivity to disturbances, From a control viewpoint it makes sense to impose a given cable configuration at the pose by setting the lengths of slack cables to larger values than the one required for the pose. To determine the best cable configuration we have proposed several ranking index [12].

Even more complex kinematic problems are involved if we assume that the cable are deformable (e.g. are elastic or catenary-like). The cable model is included in the kinematic equations for getting a complete model. We have been interested in the catenary-like model that involves inverse hyperbolic functions and is valid for steel cable of relatively high length. As the IK has never been addressed with such a model we have proposed a solving algorithm [10] that has shown that the IK may have multiple solutions but also may have no solution for poses that are reachable with non-deformable cables. In the same way the DK has several solutions [13]. Finally efficient cables interference detection for sagging cables and the management of modular CDPR, whose geometry may be changed according to the task at hand, have been addressed [9].

7.1.1.2. Cable-Driven Parallel Robots for additive manufacturing in architecture

Participant: Yves Papegay.

Easy to deploy and to reconfigure, dynamically efficient in large workspaces even with payloads, cable-driven parallel robots are very attractive for solving displacement and positioning problems in architectural building at scale 1 and seems to be a good alternative to crane and industrial manipulators in this area.

In a collaboration with CNAM and Ecole Nationale Supérieure d'Architecture Paris-Malaquais, we worked on additive manufacturing of building based on ultra-high performance concrete and developed a CDPR as a proof of concept to power a large scale 3D-printer.

A real size industrial robot will be developed by the XtreeE start-up company.

7.1.2. Assistance

This is now the core of our activity and our work on CDPR is deeply connected to this field as they are an efficient solution for mobility assistance, a high priority for the elderly, helpers and medical community. We have presented our vision of assistance robotics in several occasions [24], [25], [23].

7.1.2.1. Assessment of elderly frailty

Participants: Karim Bakal, Jean-Pierre Merlet.

The assessment of elderly frailty is a difficult concept because it involves the physical capacities of a person and its environment (health-care services, families, funds...). We consider the assessment of upper limb capabilities by looking at the joint torques τ of the arm and the maximal force F that can be exerted by the hand, which are related by the equation

$$\tau = \mathbf{J}^T F \quad (12)$$

where \mathbf{J} is a matrix which depends only upon the configuration of the arm. These equations constitute an underconstrained linear system. In biomechanics the torque τ is measured together with the configuration of the arm and the force F is evaluated by using the method of Chiacchio, that involves the pseudo-inverse of \mathbf{J}^T to calculate F . But there are several uncertainties that are neglected when using this method: the measurement errors on τ and on the configuration of the arm together with uncertainties on the physical parameters of the arm (such as the length of the bones). The method of Chiacchio provides one of the possible solutions of equation (2) and not necessary the one corresponding to the force at the hand. We use another approach based on interval analysis. We assume that all uncertainties may be bounded (τ is an interval vector τ_m , \mathbf{J}^T is an interval matrix) so that equation (2) become an interval linear system. Interval analysis then allows one to determine an approximation as accurate as wanted of the set F_s of all forces F that satisfy the equation and therefore this set includes the real force at the hand. Now assume that with the same arm configuration we measure the force at the hand, here again with some bounded uncertainties (i.e. F is an interval vector F_m). Here again we may use interval analysis applied on equation (2) in order to determine an interval vector τ_v for the τ that is guaranteed to include the real τ . Furthermore τ must be included in the intersection τ_i of τ_v and τ_m while F must be included in the intersection F_i of F_m and F_s . If τ_i is strictly included in τ_m , then we may compute a better approximation of F_s . Reciprocally if F_i is strictly included in F_m we will get a better τ_v . If one of these situation occurs we repeat the process until no significant improvement of F_s or τ_v is obtained. In a second step we consider that the uncertainties that lead to uncertainties in the matrix \mathbf{J}^T are constrained as we have to satisfy $\tau_v = \mathbf{J}^T F_s$. Here again we use interval analysis to determine if this constraint does not allow to reduce the size of the interval on the physical parameters in which case we may obtain a new \mathbf{J}^T that is included in the initial one. In turn this may allow to obtain better τ_v and F_s . The process stops when no improvement has been obtained for F_s , τ_v and the physical parameters.

To test this approach the right upper limb joint torque of 10 males and the force capacity at the right hand was measured by a dynamometer (Biodex III, Biodex Medical Systems) and respectively by a 6-axis load sensor during an experiment performed at HandiBio laboratory. The configuration of the upper limb was measured with a motion capture system (Qualisys, Sweden). The approach is currently being evaluated.

7.1.2.2. Walking analysis and Rehabilitation

Participants: Claire Maillard, Ting Wang, Jean-Pierre Merlet [correspondant].

The walkers of the ANG family allow one to determine accurately the trajectory of the walker and therefore to analyze the walking of the user. But these walkers may also be used to assess a rehabilitation process or the progress of an end-user involved in rehabilitation. For that purpose after having identified needs and requirements [17] we developed a new walker ANG-med that used infra-red distance sensors to measure the position of the subject during a rehabilitation exercise. Furthermore the software of this walker has been designed to support a message-passing scheme based on the HOP language of the INDES project team so that the walker may exchange information and control order with an external computer, together with allowing the download of new rehabilitation exercise through the robotics RAPP-store [26]. New exercises are designed as a set of such messages, that may include the calculation of exercise assessment indicators. ANG-med supports various modes: stand-alone (no external connection), passive mode (the walker only report indicator and status using a wifi connection) or full external control (an external computer fully control the walker except for emergency and real-time procedures).

ANG-med has been tested for one month in Centre Héliomarin de Vallauris and is now deployed in the rehabilitation center of MATIA in Spain, as part of the RAPP project. A start-up plan was proposed in November 2014 to transfer the walking analysis technology of HEPHAISTOS with the ANG walker in a company called Euthenia 9.2.1.3 .

7.1.2.3. Design and evaluation of assistive devices, ethics

Participants: Marc Beninati, Bernard Senach [correspondant], Jean-Pierre Merlet.

Providing appropriate support, services and information to the elderly, to their caregivers and to the medical profession, through a fleet of communicating devices must rely on a structured processes. A generic design and evaluation framework is being elaborated and will be validated through field experiments [20], [19], [18].

Assistance robotics raises many ethical questions. We started reflection about conducting experiments with frail and old people. A listing of questions to be addressed at each step of an experiment has been written (internal document). We have also hired a joint PhD student with University Bologna about the legal aspects of assistance robotics and we have initiated, together with Nathalie <Nevejans from University of Douai, a meeting with the OPECST at the French National Assembly to discuss legal and ethical aspects of robotics.

7.1.2.4. Smart Environment for Human Behaviour Recognition

Participants: Aurélien Masseur, Yves Papegay, Odile Pourtallier.

Both economic motivations due to demographic evolution and willingness of people to live independently at home when aging, facing physical impairment or recovering from injuries has raised the need for activity monitoring at home, in rehabilitation center or in retirement home. Monitoring systems provide information that can range from a broad measure of the daily activity to a precise analysis of the ability of a person performing a task (cooking, dressing, ...) and its evolution.

The broad range of needs and contexts, together with the large variety of available sensors implies the necessity to carefully think the design of the monitoring system. An appropriate system should be inexpensive and forgettable for the monitored person, should respect privacy but collect necessary data, and should easily adapt to stick to new needs. We aim to provide an assisting tool for designing appropriate monitoring systems.

As part of a PhD work, optimal motion planning of a mobile robot with range sensors to locate targets in a room has been studied. Work in progress also include algorithms to deploy infra-red barriers in a large area with several interest places, to be able to locate people. An experimental set-up is in use in the lab and data analysis methods are developed to infer people behaviors.

7.2. Miscellaneous results

7.2.1. Symbolic tools for modeling and simulation

Participant: Yves Papegay.

This activity is the main part of a long-term ongoing collaboration with Airbus whose goal is to directly translate the conceptual work of aeronautics engineers into digital simulators to accelerate aircraft design.

An extensive modeling and simulation platform has been designed which includes a dedicated modeling language for the description of aircraft dynamics models in term of formulae and algorithms, and a symbolic compiler producing as target an efficient numerical simulation code ready to be plugged into a flight simulator, as well as a formatted documentation compliant with industrial requirements of corporate memory [14].

Technology demonstrated by our prototype has been transferred, final version of our modeling and simulation environment has been delivered to Airbus in November 2012. Developer level know-how has been transferred in 2013 to a software company in charge of industrialization and maintenance of the modeling and simulation environment.

Since 2014, we are working on several enhancements and extension of functionalities, namely to enhance the performances and the numerical quality of the generated C simulation code, and ease the integration of our environment into the airbus toolbox.

HYBRID Project-Team

7. New Results

7.1. 3D User Interfaces

7.1.1. Novel 3D Interactive Techniques

THING: Introducing a Tablet-based Interaction Technique for Controlling 3D Hand Models Merwan Achibet, Anatole Lécuyer and Maud Marchal

The hands of virtual characters are highly complex 3D models that can be tedious and time-consuming to animate with current methods. We introduced the *THING* [17], a novel tablet-based approach that leverages multi-touch interaction for a quick and precise control of a 3D hand's pose [2]. The flexion/extension and abduction/adduction of the virtual fingers can be controlled for each finger individually or for several fingers in parallel through sliding motions on the surface of the tablet. We designed two variants of *THING*: (1) *MobileTHING*, which maps the spatial location and orientation of the tablet to that of the virtual hand, and (2) *DesktopTHING*, which combines multi-touch controls of fingers with traditional mouse controls for the global position and orientation of the hand model. We compared the usability of *THING* against mouse-only controls and a data glove in two controlled experiments. Results show that *DesktopTHING* was significantly preferred by users while providing performance similar to data gloves. Together, these results could pave the way to the introduction of novel hybrid user interfaces based on tablets and computer mice in future animation pipelines. This work was done in collaboration with Géry Casiez (Inria team MJOLNIR).



Figure 2. *THING* enables the control of 3D hand models (in blue) by sliding fingers along sliders arranged in a morphologically-consistent pattern on the tablet's screen. This creates a strong correspondence between user's input and pose of the controlled hand. Here, the user closes the virtual hand and then points the index finger.

Plasticity for 3D User Interfaces: New Models for Devices and Interaction Techniques Jérémy Lachoche and Bruno Arnaldi

We have introduced new models for device and interaction techniques to overcome plasticity limitations in Virtual Reality (VR) and Augmented Reality (AR) [26]. We aimed to provide developers with solutions to use and create interaction techniques that fit to the 3D application tasks and to the input and output devices available. The device model describes input and output devices and includes capabilities, limitations and representations in the real world. We also propose a new way to develop interaction techniques with an approach based on PAC and ARCH models [43]. These techniques are implemented independently from the specific devices used thanks to the proposed device model. Moreover, our approach aims to facilitate the portability of interaction techniques over different target OS and 3D frameworks. This work was done in collaboration with Thierry Duval (Lab-STICC), Éric Maisel (ENIB) and Jérôme Royan (IRT B-Com).

Dealing with Frame Cancellation for Stereoscopic Displays in 3D User Interfaces Jérémy Lacoche, Morgan Le Chénéchal, Valérie Gouranton and Bruno Arnaldi

We explored new methods to reduce ocular discomfort when interacting with stereoscopic content, focusing on frame cancellation [27]. Frame cancellation appears when a virtual object in negative parallax (front of the screen) is clipped by the screen edges; stereopsis cue lets observers perceive the object popping-out from the screen while occlusion cue provides observers with an opposite signal. Such a situation is not possible in the real world. This explains some visual discomfort for observers and leads to a poor depth perception of the virtual scene. This issue is directly linked to the physical limitations of the display size that may not cover the entire field of view of the observer. To deal with these physical constraints we introduce two new methods in the context of interactive applications. The first method consists in two new rendering effects based on progressive transparency that aim to preserve the popping-out effect of the stereo. The second method focuses on adapting the interaction of the user, not allowing him to place virtual objects in an area subject to frame cancellation. This work was done in collaboration with Sébastien Chalmé (IRT B-Com), Thierry Duval (Lab-STICC) and Éric Maisel (ENIB).

7.1.2. *Understanding Human Perception in VR*

Distance Estimation in Large Immersive Projection Systems, Revisited Ferran Argelaguet and Anatole Lécuyer

When walking within an immersive projection environment, accommodation distance, parallax and angular resolution vary according to the distance between the user and the projection walls which can influence spatial perception. As CAVE-like virtual environments get bigger, accurate spatial perception within the projection setup becomes increasingly important for application domains that require the user to be able to naturally explore a virtual environment by moving through the physical interaction space. In this work we performed two experiments which analyze how distance estimation is biased when accommodation distance, parallax and angular resolution vary [23]. The experiments were conducted in a large immersive projection setup with up to ten meter interaction range. The results showed that both accommodation distance and parallax have a strong asymmetric effect on distance judgments. We found an increased distance underestimation for positive parallax conditions as the accommodation-convergence difference increased. In contrast, we found less distance overestimation for negative and zero parallax conditions. Our findings also showed that angular resolution has a negligible effect on distance estimation. This work was done in collaboration with Anne-Hélène Olivier (MIMETIC) and Gerd Bruder (University of Hamburg).

Virtual Proxemics: Locomotion in the Presence of Obstacles in Large Immersive Projection Environments Ferran Argelaguet, Anatole Lécuyer

In the real world we navigate with ease by walking in the presence of obstacles, we develop avoidance strategies and behaviors which govern the way we locomote in the proximity of physical objects and other persons during everyday tasks. With the advances of virtual reality technology, it becomes important to gain an understanding of how these behaviors are affected in a virtual reality application. In this work, we analyzed the walking and collision avoidance behavior when avoiding real and virtual static obstacles [19]. In order to generalize our study, we considered both anthropomorphic and inanimate objects, each having his virtual and real counterpart. The results showed that users exhibit different locomotion behaviors in the presence of real and virtual obstacles, and in the presence of anthropomorphic and inanimate objects. Precisely, the results showed a decrease of walking speed as well as an increase of the clearance distance (i. e., the minimal distance between the walker and the obstacle) when facing virtual obstacles compared to real ones. Moreover, our results suggest that users act differently due to their perception of the obstacle: users keep more distance when the obstacle is anthropomorphic compared to an inanimate object and when the orientation of anthropomorphic obstacle is from the profile compared to a front position. We discussed implications on future large shared immersive projection spaces. This work was done in collaboration with Anne-Hélène Olivier (MIMETIC), Julien Pettré (MIMETIC) and Gerd Bruder (University of Hamburg).

7.1.3. *Sports and Virtual Reality*

A Methodology for Introducing Competitive Anxiety and Pressure in VR Sports Training Ferran Argelaguet and Anatole Lécuyer

Athletes' performance is influenced by internal and external factors, including their psychological state and environmental factors, especially during competition. As a consequence, current training programs include stress management. In this work, we explored whether highly immersive systems can be used for such training programs [11]. First, we proposed methodological guidelines to design sport training scenarios both on considering the elements that a training routine must have, and how external factors might influence the participant. The proposed guidelines are based on flow and social-evaluative threat theories. Second, to illustrate and validate our methodology, we designed an experiment reproducing a 10m Olympic pistol shooting competition 3. We analyzed whether changes in the environment are able to induce changes in user performance, physiological responses and the subjective perception of the task. The simulation included stressors in order to raise a social-evaluative threat, such as aggressive public behavior or unforced errors, increasing the pressure while performing the task. The results showed significant differences in the user behavior and in their subjective impressions, trends in the physiological data were also observed. Taken together our results suggest that highly immersive systems could be further used for training systems in sports. This work was done in collaboration with Frank Multon (MIMETIC).



Figure 3. The proposed methodology was illustrated and evaluated in a virtual Olympic shooting experiment. The experiment was conducted in a wide immersive projection system being able to enclose a ten meter wide shooting range with six virtual opponents and one participant.

7.1.4. Experiencing the Past in Virtual Reality

An Immersive Virtual Sailing on the 18 th -Century Ship Le Boullongne Jean-Baptiste Barreau, Florian Nouviale and Valérie Gouranton

This work is the result of the collaboration between historians and computer scientists whose goal was the digital reconstitution of “Le Boullongne”, an 18th-century merchant ship of “La Compagnie des Indes orientale” [12]. This ship has now disappeared and its reconstitution aims at understanding on-board living conditions. Three distinct research laboratories have participated in this project so far. The first, a department of naval history, worked on historical documents, especially the logbooks describing all traveling events of the ship. The second, a research laboratory in archeology, archaeoscience and history, proposed a 3D model of the ship based on the original naval architectural plans. The third, a computer science research laboratory, implemented a simulation of the ship sailing in virtual reality. This work focuses on the reconstitution of the ship in virtual reality, aiming at restoring a realistic interactive naval simulation: the 3D model of the ship has been integrated in an ocean simulation, with a physical rendering of the buoyancy. The simulation allows a user to walk around on the ship, at a scale of 1:1, and even steer it through a natural interaction. Several

characteristics of the simulation reinforce the sensation of being on-board: (1) A sonic environment mixing spatialized sounds (gulls flying, a whale swimming, wood cracking, cannons firing) and global soundscape (ocean and wind). (2) The meteorology of the simulation is dynamically modifiable; the user can increase the swell height and speed. The global illumination and wind sound vary in accordance with these parameters. The buoyancy simulation entails realistic movements of the ship. (3) Several interactions are proposed allowing the user to steer the ship with his/her hand, walk around on the ship, fire the cannons, and modify the weather. (4) Three animated sailors accompany the user in his/her sailing experience. They are wearing realistic period costumes. The immersive simulation has allowed historians to embark on “Le Boullongne” and to better understand how life was organized on-board. It has also been presented at several public exhibitions, in CAVE-like structures and HMD. This work was done in collaboration with Ronan Gagne (Univ. Rennes 1), Yann Bernard (CReAAH) and Sylviane Llinares (CERHIO, UBS Lorient).

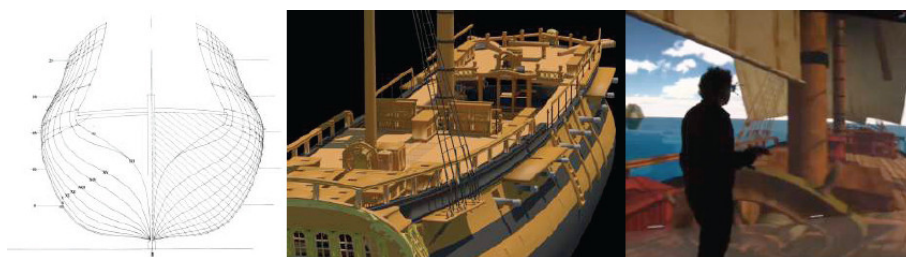


Figure 4. Digital reconstitution of “Le Boullongne”. From architectural plans to virtual reality implementation.

Touching and interacting with inaccessible cultural heritage Valérie Gouranton and Bruno Arnaldi

Sense of touch provides a particular access to our environment, enabling a tangible relation with it. In the particular use case of cultural heritage, touching the past, apart from being a universal dream, can provide essential information to analyze, understand, or restore artifacts. However, archaeological objects cannot always offer a tangible access, either because they have been destroyed or too damaged, or because they are part of a larger assembly. In other cases, it is the context of use that has become inaccessible, as it is related to an extinct activity. In [15] we proposed a workflow based on a combination of computed tomography, 3D images, and 3D printing to provide concrete access to cultural heritage, and we illustrate this workflow in different contexts of inaccessibility. These technologies are already used in cultural heritage, but seldom combined, and mostly for exceptional artifacts. We proposed to combine these technologies in case studies corresponding to relevant archaeological situations.

This work was done in collaboration with Théophile Nicolas (INRAP), Ronan Gagne (Univ. Rennes 1), Cédric Tavernier (Image ET) and Quentin Petit (CNRS).

3D reconstruction of the Loyola sugar plantation and virtual reality applications Jean-Baptiste Barreau, Valérie Gouranton

Discovered in 1988, the Loyola sugar plantation, owned by the Jesuits in French Guiana, is a major plantation of colonial history and slavery. Ongoing archaeological excavations have uncovered the Jesuit’s house and the outbuildings usually associated with a plantation such as a chapel and its cemetery, a blacksmith shop, a pottery, the remains of the entire sugar production (a windmill, a boiler and a dryer), coffee and indigo warehouses etc. Based on our findings and our network with 3D graphic designers and researchers in virtual reality, a 3D restitution integrated within a virtual reality platform was initiated to develop a better understanding of the plantation and its surrounding landscape. A specific work on the interactive changes of sunlight and animal sounds aimed to reconstruct a coherent evolution during one day of the site’s environment [21].

This work was done in collaboration with Quentin Petit (CNRS), Yann Bernard (CReAAH), Reginald Auger (Laval University, Canada), Yannick Le Roux (Laval University, French Guiana) Ronan Gagne (IMMER-SIA), and Cédric Tavernier (Image ET).

7.2. Physically-Based Simulation and Multisensory Feedback

7.2.1. Interactive Physically-Based Simulation

Aggregate constraints for virtual manipulation with soft fingers, Maud Marchal, Anthony Talvas



Figure 5. Interaction with deformable fingers generates many interconnected contact points which are expensive to solve with friction. Our approach aggregates contact constraints per phalanx with torsional friction. The subsequent increase in performance allows real time dexterous manipulation of virtual objects using soft fingers.

Interactive dexterous manipulation of virtual objects remains a complex challenge that requires both appropriate hand models and accurate physically-based simulation of interactions. In [16], we proposed an approach based on novel aggregate constraints for simulating dexterous grasping using soft fingers. Our approach aims at improving the computation of contact mechanics when many contact points are involved, by aggregating the multiple contact constraints into a minimal set of constraints. We also introduced a method for non-uniform pressure distribution over the contact surface, to adapt the response when touching sharp edges. We used the Coulomb-Contensou friction model to efficiently simulate tangential and torsional friction. We showed through different use cases that our aggregate constraint formulation is well-suited for simulating interactively dexterous manipulation of virtual objects through soft fingers, and efficiently reduces the computation time of constraint solving. This work was done in collaboration with Christian Duriez (Inria team DEFROST) and Miguel Otaduy (Univ. Rey Juan Carlos, Madrid, Spain).

7.2.2. Multimodal Feedback

Elastic-Arm: Human-scale passive feedback for augmenting interaction and perception in virtual environments Merwan Achibet, Adrien Girard, Maud Marchal, Anatole Lécuyer

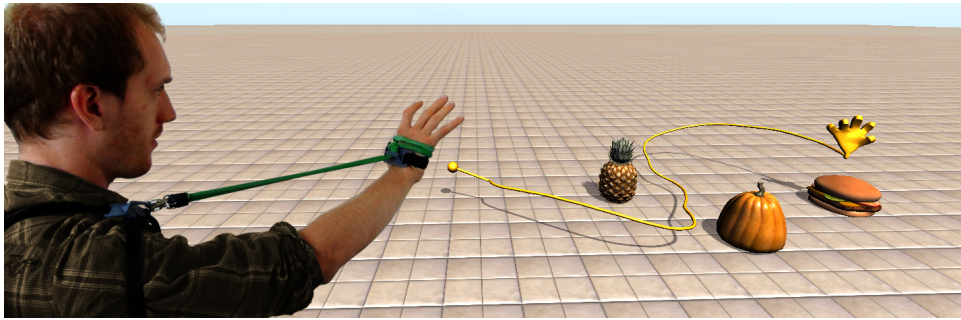


Figure 6. The Elastic-Arm is a body-mounted armature that provides egocentric passive haptic feedback. It presents an alternative to more complex active haptic devices that are generally less adapted to large immersive environments. In this example, a user performs a selection task by stretching his virtual arm using a combination of the Bubble and Go-Go techniques reimplemented with our system.

Haptic feedback is known to improve 3D interaction in virtual environments but current haptic interfaces remain complex and tailored to desktop interaction. In [18], we introduced the *ElasticArm*, a novel approach for incorporating haptic feedback in immersive virtual environments in a simple and cost-effective way. The Elastic-Arm is based on a body-mounted elastic armature that links the user's hand to her shoulder. As a result, a progressive resistance force is perceived when extending the arm. This haptic feedback can be incorporated with various 3D interaction techniques and we illustrate the possibilities offered by our system through several use cases based on well-known examples such as the *Bubble* technique, *Redirected Touching*, and pseudo-haptics. These illustrative use cases provide users with haptic feedback during selection and navigation tasks but they also enhance their perception of the virtual environment. Taken together, these examples suggest that the Elastic-Arm can be transposed in numerous applications and with various 3D interaction metaphors in which a mobile haptic feedback can be beneficial. It could also pave the way for the design of new interaction techniques based on human-scale egocentric haptic feedback.

Visual vibrations to simulate taps on different materials Maud Marchal, Anatole Lécuyer

In [40], we presented a haptic visualization technique for conveying material type through visual feedback, expressed as visible decaying sinusoidal vibration resulting from tapping an object. The technique employs cartoon-inspired visual effects and modulates the scale of the vibration to comply with visual perception. The results of a user study showed that participants could successfully perceive three types of material (rubber, wood, and aluminum) using our novel visual effect. This work was done in collaboration with Taku Hachisu and Hiroyuki Kajimoto (Univ. of Electro Communication, Tokyo, Japan).

7.2.3. GPU-based Collision Detection in Virtual Environments

GPU Ray-Traced Collision Detection: Fine Pipeline Reorganization François Lehericey, Valérie Gouranton, Bruno Arnaldi

Ray-tracing algorithms can be used to render a virtual scene and to detect collisions between objects. Numerous ray-tracing algorithms have been proposed which use data structures optimized for specific cases (rigid objects, deformable objects, etc.). Some solutions try to optimize performance by combining several algorithms to use the most efficient algorithm for each ray. In [31], we presented a ray-traced collision detection pipeline that improves the performance on a graphic processing unit (GPU) when several ray-tracing algorithms are used.

When combining several ray-tracing algorithms on a GPU, a well-known drawback is thread divergence among work-groups that can cause loss of performance by causing idle threads. We avoid branch divergence

by dividing the ray tracing into three steps with appended buffers in between. We also show that prediction can be used to avoid unnecessary synchronizations between the CPU and GPU. Applied to a narrow-phase collision detection algorithm, results show an improvement of performance up to 2.7 times.

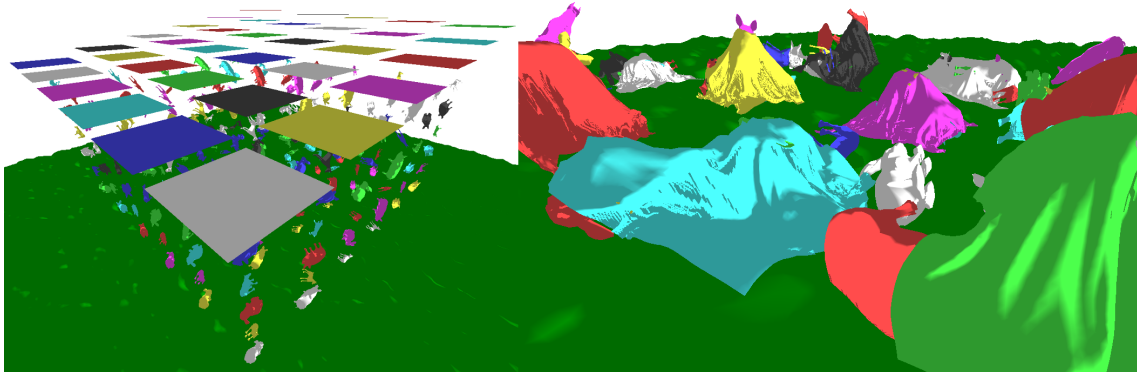


Figure 7. 216 concave objects fall on an irregular ground and 36 deformable sheets fall over them [31].

GPU Ray-Traced Collision Detection for Cloth Simulation François Lehericéy, Valérie Gouranton, Bruno Araldi

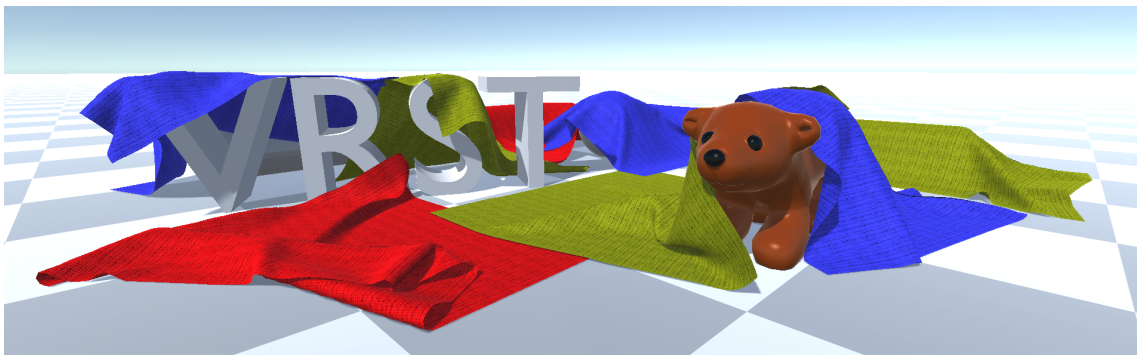


Figure 8. Our method can perform collision detection between clothes and handle self collision detection [30].

In [30], we proposed a method to perform collision detection with cloths with ray-tracing at an interactive frame-rate. Our method is able to perform collision detection between cloths and volumetric objects (rigid or deformable) as well as collision detection between cloths (including auto-collision). Our method casts rays between objects to perform collision detection, and an inversion-handling algorithm is introduced to correct errors introduced by discrete simulations. GPU computing is used to improve the performances by parallelizing the ray-tracing. Our implementation handles scenes containing deformable objects at an interactive frame-rate, with collision detection lasting a few milliseconds.

7.2.4. Medical Applications

Real-time tracking of deformable targets in 3D ultrasound images Maud Marchal

In [35], [36], we presented a novel approach for tracking a deformable anatomical target within 3D ultrasound volumes. Our method is able to estimate deformations caused by the physiological motions of the patient. The displacements of moving structures are estimated from an intensity-based approach combined with a physically-based model and has therefore the advantage to be less sensitive to the image noise. Furthermore, our method does not use any fiducial marker and has real-time capabilities. The accuracy of our method is evaluated on real data acquired from an organic phantom. The validation is performed on different types of motions comprising rigid and non-rigid motions. Thus, our approach opens novel possibilities for computer-assisted interventions where deformable organs are involved.

Our approach was also evaluated on the MICCAI CLUST'15 challenge 3D database. We achieved a mean tracking error of 1.78 mm with an average computation time of 350 ms per frame, ranking our method first during the on-site challenge [34]. This work was done in collaboration with Lucas Royer, Anthony Le Bras and Guillaume Dardenne (IRT bcom), and Alexandre Krupa (Inria team LAGADIC).

Statistical study of parameters for deep brain stimulation automatic pre-operative planning of electrodes trajectories Maud Marchal

Automatic methods for pre-operative trajectory planning of electrodes in Deep Brain Stimulation are usually based on the search for a path that resolves a set of surgical constraints to propose an optimal trajectory. In [13], we studied the use of parameters based on real trajectories of surgeons. For that purpose we firstly retrieve the actual weighting factors used by neurosurgeons thanks to a retrospective study, secondly we compare the results from two different hospitals to evaluate their similarity, and thirdly we compare these trends to the weighting factors usually empirically set in most current approaches. We proposed two approaches, one based on a stochastic sampling and the other on an exhaustive search. In each case, we get a sample of combinations of weighting factors along with a measure of their quality, i.e. the similarity between the optimal trajectory they lead to and the trajectory manually planned by the surgeon as a reference. Then visual and statistical analysis are performed on the number of occurrences and on the rank means. We performed our study on 56 retrospective cases from two different hospitals. We could observe a trend of the occurrence of each weight on the number of occurrences. We also proved that each weight had a significant influence on the ranking. Additionally, we observed no influence of the medical center parameters, suggesting that the trends were comparable in both hospitals. Finally, the obtained trends were confronted to the usual weights chosen by the community, showing some common points but also some discrepancies. These results tend to show a predominance of the choice of a trajectory close to a standard direction. Secondly, the avoidance of the vessels or sulci seems to be sought in the surroundings of the standard position. The avoidance of the ventricles seem to be less predominant, but this could be due to the already reasonable distance between the standard direction and the ventricles. The similarity of results between two medical centers tend to show that it is not an exceptional practice. This work was done in collaboration with Caroline Essert and Antonio Capobianco (Univ. Strasbourg), Claire Haegelen and Pierre Jannin (LTSI, Rennes), Sara Fernandez-Vidal, Carine Karachi and Eric Bardinet (Institut du Cerveau et de la Moëlle Epinière, Paris).

7.3. Collaborative Virtual Environments

Asymmetric Remote Collaboration in Mixed Reality: Awareness and Navigation Morgan Le Chénéchal, Valérie Gouranton and Bruno Arnaldi

We first focused on the lack of mutual awareness that may appear in many situations and we evaluated different ways to present the distant user and his actions in the Virtual Environment (VE) in order to understand his perception and cognitive process. We focused on a common case consisting in estimating accurately the time at which a distant user analyzed the meaning of a remotely pointed object. Amongst others, our experimental results presented at CTS [28], show that expertise of the users influences on how they estimate the distant activity and the type of applied strategies.

Then, in a similar asymmetric setup, we proposed a demo at IEEE VR to deal with real estate business. In this context, it is quite difficult for estate agents to make customers understand the potential and the volumes of free spaces. The demo aimed to solve these issues based on a laying out scenario in which a seller and a

customer collaborate. As the roles of both users are different, we proposed an asymmetric collaboration where the two users do not use the same interaction setup and do not benefit from the same interaction capabilities.

Last, we focused on a remote collaborative maintenance scenario in which a remote expert helps an operator in performing a physical task [9](#). Our system is based on a VR setup for the remote expert in order to virtually co-locate him in the real workspace, and an AR interface for the display of the helping gestures to the agent. In a preliminary user study, we evaluated the performance of our system in a navigation task, and we presented results at ICAT-EGVE [\[29\]](#).

This work was done in collaboration with Thierry Duval (Lab-STICC) and Jérôme Royan (IRT B-Com).



Figure 9. Remote collaborative maintenance using mixed reality.

High-Level Components for Developing Collaborative and Interactive Virtual Environments Rozenn Bouville, Valérie Gouranton, Thomas Boggini, Florian Nouviale and Bruno Arnaldi

We proposed a framework called #FIVE (Framework for Interactive Virtual Environments) for the development of interactive and collaborative virtual environments [\[22\]](#). It has been developed for an easier and a faster design and development of virtual reality applications. It was designed with a constant focus on re-usability with as few hypotheses as possible on the final application in which it could be used. Whatever the chosen implementation for the Virtual Environment (VE), #FIVE : (1) provides a toolkit that eases the declaration of possible actions and behaviours of objects in the VE, (2) provides a toolkit that facilitates the setting and the management of collaborative interactions in a VE, (3) is compliant with distribution of the VE on different setups and (4) proposes guidelines to efficiently create a collaborative and interactive VE. It is composed of several modules, among them, two core modules : the relation engine and the collaborative interaction engine. On the one hand, the relation engine manages the relations between the objects of the environment. On the other hand, the collaborative interaction engine manages how users can collaboratively control objects. The modules that compose the #FIVE framework can be used either independently or simultaneously, depending on the requirements of the application. They can also communicate and work with other modules thanks to an API. For instance, a scenario engine can be plugged to any or both of the #FIVE modules if the application is scenario-based. #FIVE has already been used in VR applications by several members of our team (see section [6.5](#)). The feedbacks are rather positive and we intend to further develop #FIVE with additional functionalities, notably by extending it to the control of avatars whether they are controlled by a user or by the system.

High-Level Components for Developing Collaborative Scenarios Guillaume Claude, Valérie Gouranton and Bruno Arnaldi

We were interested in the description of activities of actors in Collaborative Virtual Environments for Training to team working on procedures. We have proposed #SEVEN, a model for the description of procedures as

Collaborative Virtual Environments Scenarios (see also section 6.6). In [25] we have demonstrated the abilities of this model to be adapted to a wide range of use cases. We showed that it can adapt its abstraction level to the required guidance level and describe more or less complex unfolding of events. In [24] we have provided a novel approach to the distribution of the actions between the actors of the simulation by using an action filtering model in conjunction with a reactive team model. The action filtering model uses data about the actors such as their abilities or their rights. Our reactive team model can be used to define relationships between the team members and the effects of inner rules of the team upon the involvement of the actors in the procedure. To our knowledge, our solution is the closest to the existing models proposed by the social science domain known as role theory. Our work has been applied to several domains, including the training of scrub nurses to neurosurgery procedures 10 .



Figure 10. The #FIVE and #SEVEN models used in the S3PM project to provide an interactive environment and define collaborative scenarios and handle dynamic team structures in a surgical context.

7.4. Brain-Computer Interfaces

7.4.1. Novel Usages of BCI

Mind-Window: Real-Time Brain Activity Visualization Using Tablet-Based Augmented Reality and EEG for Single or Multiple Users, Anatole Lécuyer, Jonathan Mercier, Maud Marchal



Figure 11. Our novel “Mind-Window” approach enables one or multiple users to visualize the brain activity of a person in real-time by using tablets and augmented reality. It proposes to see through the tablet a virtual brain model “as if the skull is transparent”. The display of the virtual brain is updated in real-time according to the real brain activity of the person which is measured thanks to an EEG headset.

We introduced a novel approach, called the “Mind-Window”, for real-time visualization of brain activity [33]. The Mind-Window enables one or multiple users to visualize the brain activity of another person as if his/her skull was transparent. Our approach relies on the use of multiple tablet PCs that the observers can move around the head of the observed person wearing an electroencephalography cap (EEG). A 3D virtual brain model is superimposed to the head of the observed person using augmented reality by tracking a 3D marker placed on top the head. The EEG cap records the electrical fields emitted by the brain, and they are processed in real-time to update the display of the virtual brain model. Several visualization techniques are proposed such as an interactive cutting plane which can be manipulated with touch-based inputs on the tablet. The Mind-Window approach could be used for medical applications, e.g. by providing a simple way for physicians to diagnose and observe brain activity of patients. Teachers could also use our system to teach brain anatomy/activity and EEG features, e.g., electrodes localization, electrical patterns, etc. Finally, video conferences or video games could be “brain-augmented”, making use of the Mind-Window for entertainment purposes.

B-C-Invisibility Power: Optical Camouflage Based on Mental Activity in Augmented Reality, Anatole Lécuyer, Jonathan Mercier, Maud Marchal



Figure 12. The “B-C-Invisibility power” enables users to become virtually invisible by performing mental tasks. Brain signals are extracted using EEG electrodes and analyzed within the BCI.

In the context of the ANR project HOMO-TEXTILUS which focuses on the design of novel “smart clothes”, we introduced a kind of “invisibility cloak”: an interactive approach for using Brain-Computer Interfaces for controlling optical camouflage called “B-C-Invisibility power”. We proposed to combine augmented reality and BCI technologies to design a system which somehow provides the “power of becoming invisible” [32]. Our optical camouflage is obtained on a PC monitor combined with an optical tracking system. A cut out image of the user is computed from a live video stream and superimposed to the prerecorded background image using a transparency effect. The transparency level is controlled by the output of a BCI, making the user able to control her invisibility directly with mental activity. The mental task required to increase/decrease the invisibility is related to a concentration/relaxation state. Results from a preliminary study based on a simple video-game inspired by the Harry Potter universe could notably show that, compared to a standard control made with a keyboard, controlling the optical camouflage directly with the BCI could enhance the user experience and the feeling of “having a super-power”.

7.4.2. BCI Methodology and Techniques

A methodological framework for applications combining BCI and videogames, Anatole Lécuyer

We have proposed a user-centered methodological framework [41] to guide design and evaluation of applications based on Brain-Computer Interface (BCI). Our framework is based on the contributions of ergonomics

to ensure that these applications are well suited for end-users. It provides methods, criteria and metrics to perform the phases of the human-centered design process aiming to understand the context of use, specify the user needs and evaluate the solutions in order to define design choices. Several ergonomic methods (e.g., interviews, longitudinal studies, user based testing), objective metrics (e.g., task success, number of errors) and subjective metrics (e.g., mark assigned to an item) are suggested to define and measure the usefulness, usability, acceptability, hedonic qualities, appealingness, emotions related to user experience, immersion and presence to be respected. The benefits and contributions of our user centred framework for the ergonomic design of videogames based on BCI were also discussed.

This work was done in collaboration with Fabien Lotte (Inria team POTIOC).

Feasibility and specificity of simultaneous EEG and fMRI, Marsel Mano, Lorraine Perronnet, Jussi Lindgren, Anatole Lécuyer

In the field of fMRI, Arterial Spin Labeling (ASL) imaging relies on control and label radio-frequency pulses. This generates alternate gradient patterns as well as higher specific absorption rate (SAR). To date, only a few studies have addressed the issue of connecting EEG signal to ASL perfusion. Furthermore, previous studies have shown reduced blood-oxygen-level dependent (BOLD) signal-to-noise ratio (SNR) in the presence of EEG. ASL being a low SNR technique, the aim of this study was to assess ASL-EEG at 3T in terms of safety as well as EEG and magnetic resonance signal quality. Our experimental results show that ASL-EEG can be safely performed [20] [38]. Standard ASL acquisitions generated more than 2.5-fold SAR increase compared to a standard BOLD echo planar imaging sequence. This corresponded to up to 4°C temperature increase on the bundle, yet not exceeding 36°C. Gradient artifact correction of the EEG signal by average artifact subtraction was generally good for BOLD-EEG and ASL-EEG. However, residual gradient artifacts affecting 1% of the pulsed ASL-EEG data have to be considered. Further research is needed to understand the artifact variability and to develop an appropriate correction strategy. No residual artifacts were observed for alternating control and label pulses ASL-EEG. Neither a change of the number of reference volumes for artifact subtraction nor an independent component analysis could help tackle this gradient artifact correction issue. Regarding magnetic resonance imaging, a 20% loss in SNR was observed when compared to acquisitions performed without EEG. Taken together our results suggest that EEG and ASL MRI can be simultaneously combined for the purpose of real-time experiments which could for instance be envisioned in our HEMISFER project.

This work was done in collaboration with VISAGES team.

ILDA Team

7. New Results

7.1. An Evaluation of Interactive Map Comparison Techniques

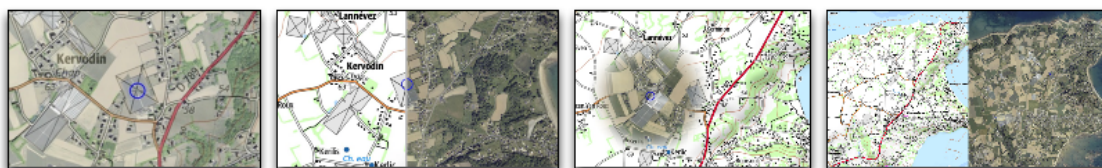


Figure 4. Empirical evaluation of multiplexing strategies using juxtaposition or overlaying for spatially-registered map comparison tasks [4]. Research conducted in the context of ANR project MapMuxing with IGN (Institut National de l'Information Géographique et Forestière).

Geovisualization applications typically organize data into layers. These layers hold different types of geographical features, describe different characteristics of the same features, or represent those features at different points in time. Layers can be composited in various ways, most often employing a juxtaposition or superimposition strategy, to produce maps that users can explore interactively. From an HCI perspective, one of the main challenges is to design interactive compositions that optimize the legibility of the resulting map and that ease layer comparison. We characterized five representative techniques, and empirically evaluated them using a set of real-world maps in which we purposefully introduced six types of differences amenable to inter-layer visual comparison. We discussed the merits of these techniques in terms of visual interference, user attention and scanning strategy. Those results can help inform the design of map-based visualizations for supporting geo-analysis tasks in many application areas.

This work was published at ACM CHI 2015 [4], and received an honorable mention (top 5% of all submissions).

7.2. Reciprocal Drag and Drop

Drag-and-drop has become ubiquitous, both on desktop computers and touch-sensitive surfaces. It is used to move and edit the geometry of elements in graphics editors, to adjust parameters using controllers such as sliders, or to manage views (e.g., moving and resizing windows, panning maps). Reverting changes made via a drag-and-drop usually entails performing the reciprocal drag-and-drop action. This can be costly as users have to remember the previous position of the object and put it back precisely. We introduced the DnD^{-1} model that handles all past locations of graphical objects. We redesigned the Dwell-and-Spring widget to interact with this history. Applications can implement DnD^{-1} to enable users to perform reciprocal drag-and-drop to any past location for both individual objects and groups of objects. We performed two user studies, whose results show that users understand DnD^{-1} , and that Dwell-and-Spring enables them to interact with this model effectively.

This work was published in ACM ToCHI [1].

7.3. SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays

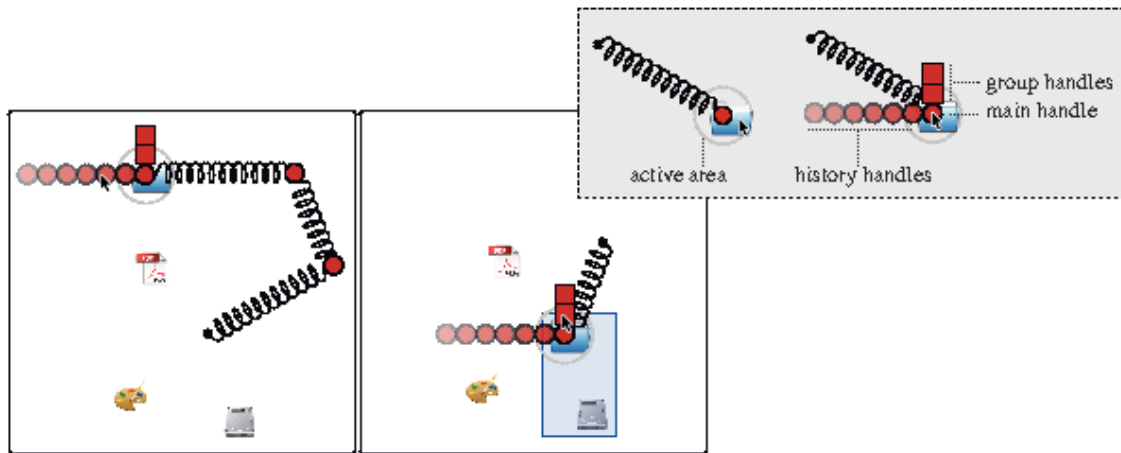


Figure 5. Navigating a graphical object's direct manipulation history as captured by the DnD^{-1} model, using the Dwell-and-Spring widget.

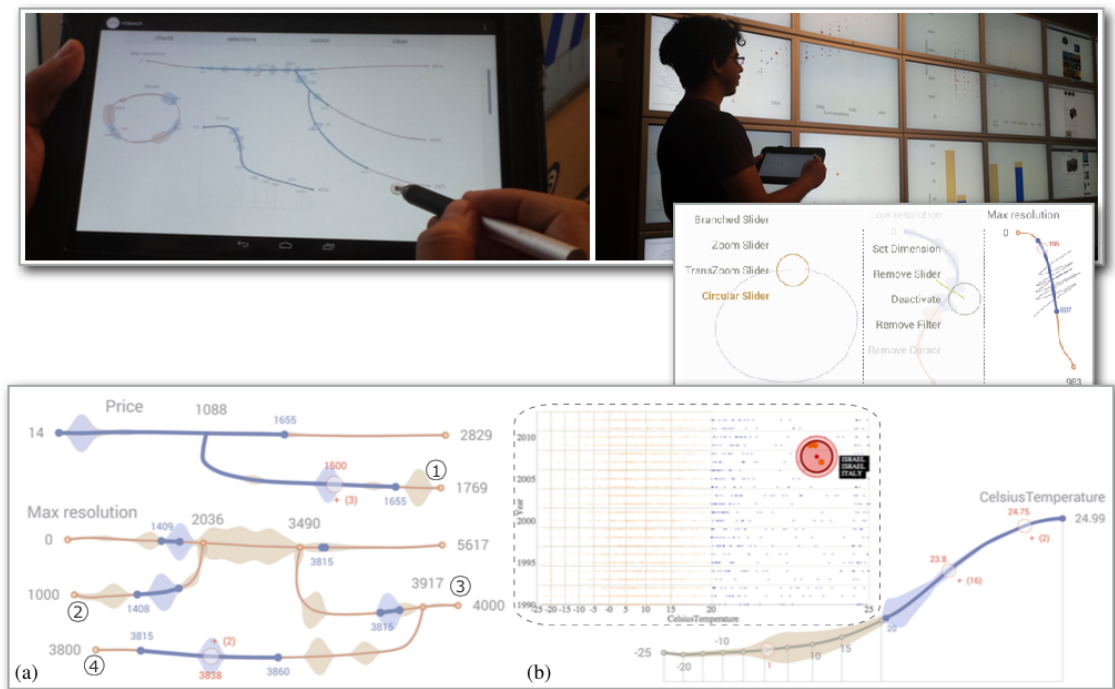


Figure 6. (top) The user sketching their sliders on the fly (left), to interact with their data on the wall display (right). Menus and simple gestures (middle) are enough to create complex sliders (bottom) that can help explore data at different granularities.

Given our interest in how to effectively interact with wall displays, we have started investigating ways to empower end users, by allowing them to easily create themselves their interfaces. We introduced a sketching interface that runs on mobile devices, and allows users to explore multi-dimensional datasets on wall displays by sketching on the fly the interactive controllers they require. We demonstrated this concept with *SketchSliders*, range sliders that users can freely sketch on the mobile surface to customize their exploration. A small combination of sketches and gestures allows the creation of complex interactive sliders, such as circular sliders for periodic data, slider branches for detailed interaction, and fisheye transformation sliders. We augmented sliders with a suite of tools, such as markers, slider cursors, and approximate views of data distributions. These designs were inspired by a design study with three visualization experts, and validated through a user study with six experts using our system.

This work was published at ACM CHI 2015 [9], and received an honorable mention (top 5% of all submissions).

7.4. Ultra-high-resolution Wall-sized Displays

We have worked on the following other projects, also related to the interactive visualization of large datasets on ultra-high-resolution wall displays:

- Mid-air Pointing on Ultra-Walls [5]. The size and resolution of ultra-high resolution wall-sized displays (“ultra-walls”) make traditional pointing techniques inadequate for precision pointing. We studied mid-air pointing techniques that can be combined with other, domain-specific interactions. We explored the limits of existing single-mode remote pointing techniques and demonstrated theoretically that they do not support high-precision pointing on ultra-walls. We then explored solutions to improve mid-air pointing efficiency: a tunable acceleration function and a framework for dual-precision (DP) techniques, both with precise tuning guidelines.
- WallTweet: A Knowledge Ecosystem for Supporting Situation Awareness [20]. Tweets are an important source of information during large-scale events, like tornados or terrorist attacks. Yet, tweets are hard to visualize and put in a geographical context: large quantities of tweets get sent in a short period, that vary greatly in content and relevance with respect to the crisis at hand. WallTweet is a tweet visualization designed for wall displays and aimed at improving the situation awareness of users monitoring a crisis event utilizing tweets.
- The monitoring of road traffic data on wall-sized displays [15]. Road traffic is a complex system that can be very unstable. A little perturbation can lead to a traffic-crippling congestion. To avoid such situations, researchers attempt to model traffic in order to prevent congestions and optimize traffic flow. Traffic is also continually monitored by operators in traffic control rooms. We designed an interactive system to monitor traffic on a wall display, that is coupled to traffic modeling algorithms. The system enables users to interactively adjust traffic parameter settings and visualize the impact of these adjustments at both a local and global scale.

IMAGINE Project-Team

6. New Results

6.1. User-centered Models for Shapes and Shape Assemblies

- **Scientist in charge:** Stefanie Hahmann.
- **Other permanent researchers:** Marie-Paule Cani, Jean-Claude Léon, Damien Rohmer.

Our goal, is to develop responsive shape models, i.e. 3D models that respond in the expected way under any user action, by maintaining specific application-dependent constraints (such as a volumetric objects keeping their volume when bent, or cloth-like surfaces remaining developable during deformation, etc). We are extending this approach to composite objects made of distributions and/or combination of sub-shapes of various dimensions.

6.1.1. Developable Surfaces

Participants: Antoine Begault, Marie-Paule Cani, Stefanie Hahmann, Damien Rohmer, Camille Schreck.



Figure 5. Example of developable model generation from sketches in [10] and for virtual paper interaction [14].

Developable surfaces are surfaces which can be unflattened on a plane without being stretched nor squeezed. In other words, they can be made from 2D pattern without change of lengths. They are usually hard to model efficiently as the length condition is non linear. We developed this year two different applications for developable surfaces, once applied for leather product designer, and the other one to virtual paper deformation.

We developed a method to generate 3D models for garments and leather products from designer sketches. Given two or three orthogonal sketched views depicting the silhouette, the seams, and the folds, we automatically compute a 3D developable surface and the corresponding 2D patterns which fits the silhouette and exhibits the designed folds. Our method can handle complex cases where the 2D silhouette actually correspond to a non planar and discontinuous curve on the 3D surface. We also proposed a new efficient approach to improves the developability of the resulting surface while preserving the pre-designed folds. This work has been published in ACM Transactions on Graphics [10], and we presented it in SIGGRAPH Asia in November.

Within the PhD work of Camille Shreck, we developed the first interactive 3D virtual model of crumpled paper. Deforming virtual paper is especially challenging to model efficiently has crumpling can be seen as singularities on the surface, leading therefore to non smooth surfaces which do not fit well to standard physically based deformation model. We proposed in this work a new geometrical representation of surface especially adapted to model non smooth developable surfaces as a set of planes, cylinders, and generalized cones meeting at the discontinuities of the surface. Our model can dynamically adapts to the surface deformation and to new crumples, while been associated to an optimal mesh triangulation containing very few triangles. Our interactive deformation model interleaves a standard Finite Element Model on the coarse

triangular mesh to guide the general deformation, with a geometrical steps adapting our surface structure to optimally sample the degrees of freedom of the crumpled paper. This work as been accepted for publication in ACM Transaction on Graphics [14], has been presented at the conference WomEncourage [29], and as a communication in AFIG [33].

6.1.2. Procedural models for shape assemblies

Participants: Marie-Paule Cani, Damien Rohmer, Ulysse Vimont.

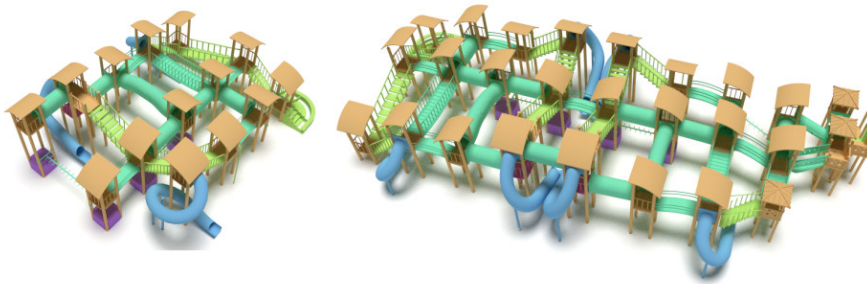


Figure 6. Example of shape assemblies before and after subpart-replacement in [11].

A popular mode of shape synthesis involves mixing and matching parts from different objects to form a coherent whole. In collaboration with the University College London, Universiteit Utrecht, and KAUST, we proposed a method to automatically detect replaceable subparts within a complex assembly. In this work, we model the geometrical assembly as a graph where each node represent a single component, and the edges represents inter-part connectivity. Our method analyses this graph to detect similar inter-part connectivity enabling to exchange or mix sub-structures to synthesise new geometrical models. This work has been published in Eurographics [11].

6.1.3. Toward Functional CAD assemblies

Participants: Pablo Covès, Harold Vilmart, Robin Roussel, Damien Rohmer, Marie-Paule Cani, Jean-Claude Léon.

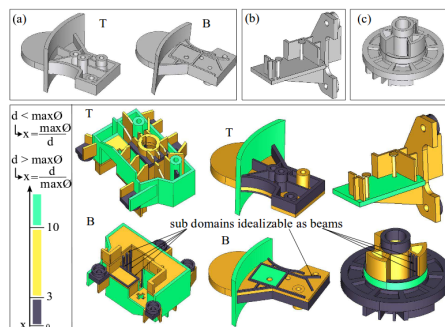


Figure 7. Example of shape idealization in [5].

We chose to focus on man-made objects to tackle the topic of shape assemblies. This is two-folds since CAD models of virtual industrial prototypes provide an excellent, real-size test-bed for our methods. Moreover, this is perfectly fitting the demand from industrial partners such as EDF and Airbus Group. On a complementary basis, we have initiated a partnership with UCL (University College London) to address function-preserving assembly deformation.

Assemblies representing products are most often reduced to a collection of independent CAD models representing each component. To our knowledge, there has been no approach proposed to generate CAD assembly models from 3D scans. An approach is initiated with a partnership with LIRIS (R. Chaine and J. Digne) and EDF in the framework of a Rhône-Alpes region project (Potasse) starting with the PhD of P. Coves.

Following the work of [43],[5], [38], partnership with Inria GRAPHIK team (F. Ulliana) has been set up and a deductive logic framework has been coupled to the SALOME platform with the insertion of an ontology describing a subset of a product structure. This partnership is developed with the internship of H. Vilmart to evolve toward an intrinsic, knowledge-based representation of a product structure that takes into account the isometries of components using our prior work about symmetry analyses [42]. The description of components through this product structure aims at supporting the generation of CAD assembly models from 3D scans to be able to derive functionally meaningful constraints of relative positions of components extracted from scans.

In the scope of the ERC Expressive, a partnership has been set up with N. Mitra (UCL) with the starting PhD of R. Roussel addressing function-preserving assembly deformation.

6.2. Motion & Sound Synthesis

- **Scientist in charge:** François Faure.
- **Other permanent researchers:** Marie-Paule Cani, Damien Rohmer, Rémi Ronfard.

Animating objects in real-time is mandatory to enable user interaction during motion design. Physically-based models, an excellent paradigm for generating motions that a human user would expect, tend to lack efficiency for complex shapes due to their use of low-level geometry (such as fine meshes). Our goal is therefore two-folds: first, develop efficient physically-based models and collision processing methods for arbitrary passive objects, by decoupling deformations from the possibly complex, geometric representation; second, study the combination of animation models with geometric responsive shapes, enabling the animation of complex constrained shapes in real-time. The last goal is to start developing coarse to fine animation models for virtual creatures, towards easier authoring of character animation for our work on narrative design.

6.2.1. Real-time physically-based models

Participants: Marie-Paule Cani, Francois Faure, Pierre-Luc Manteaux, Richard Malgat, Matthieu Nesme.

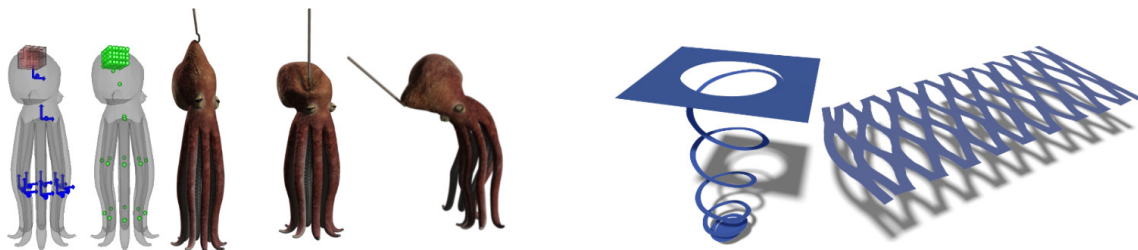


Figure 8. Left: Mixing a coarse frame-based simulation to a local FEM patch from [24]. Right: Frame based simulation for surface cutting in [25].

We keep on improving fundamental tools in physical simulation, such as new insight on constrained dynamics [15] at Siggraph. This allows more stable simulations of thin inextensible objects. A new extension of our volumetric contact approach (Siggraph 2010 and 2012) has been proposed [17] to apply rotational reaction to contact according to the shape of the contact area.

We have proposed an original approach to multi-resolution simulation, in which arbitrary deformation fields at different scales can be combined in a physically sound way[24]. This contrasts with the refinement of a given technique, such as hierarchical splines or adaptive meshes.

Following the success of frame-based elastic models (Siggraph 2011), a real-time animation framework provided in SOFA and currently used in many of our applications with external partners, we proposed an extension to the cutting of surface objects [25], in collaboration with Berkeley, where Pierre-Luc Manteaux spent 4 months at the end of 2014.

6.2.2. Simulating paper material with sound

Participants: Marie-Paule Cani, Pierre-Luc Manteaux, Damien Rohmer, Camille Schreck.



Figure 9. Left: Example of our paper tearing model in [23]. Right: Our sound synthesis for paper crumpling in [34].

Extending our results on animating paper crumpling, we proposed to synthesise the sound associated to paper material. We proposed a real time model dedicated to paper tearing. In this work, we model the specific case when two hands are tearing a flat sheet of paper on a table, in this case we synthesise procedurally the geometrical deformation of the sheet using conical surface, the tearing using a procedural noise map, and the tearing sound as a modified white noise depending on the speed of action. This work has been published in Motion in Games conference [23].

We are also developing a sound synthesis method for paper crumpling. The geometrical surface deformation is analysed to drive a procedurally synthesized friction sound and a data driven crumpling sound. We are currently developing this work and did a first communication to AFIG conference [34].

6.2.3. Animating anatomy

Participants: Armelle Bauer, Ali Hamadi Dicko, Francois Faure, Olivier Palombi, Damien Rohmer.

A real-time spine simulation model leveraging the multi-model capabilities of SOFA was presented in an international conference on biomechanics [7]. We also used a biomechanical model to regularize real-time motion capture and display, and performed live demos at the Emerging Technologies show of Siggraph Asia, Kobe, Japan [41].

We are developing an ontology-based virtual human embryo development model. In one side, a dedicated ontology stores the anatomical knowledge about organs' geometry, relations, and development rules. On the other side, we synthesize an animated visual 3D model using the informations of the ontology. This work can be seen as a first step toward interactive development anatomy teaching, or simulation, based on an ontology storing existing medical knowledge. This work has been published in the Journal of Biomedical Semantics [12].

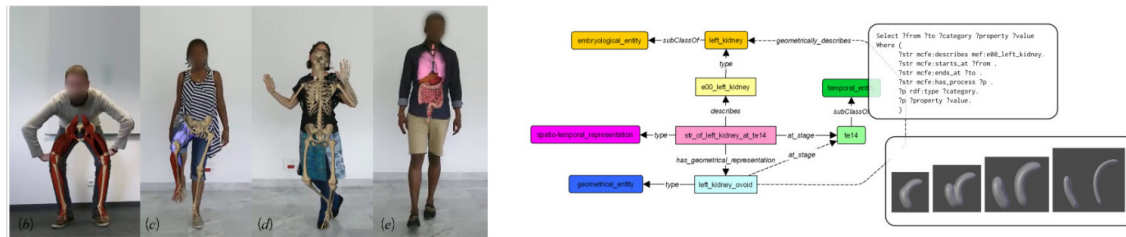


Figure 10. Left: Example of the Living Book of Anatomy in [41]. Right: Example of kidney development ontology and its 3D modeling in [12].

6.3. Knowledge-based Models for Narrative Design

- **Scientist in charge:** Rémi Ronfard.
- **Other permanent researchers:** Marie-Paule Cani, François Faure, Jean-Claude Léon, Olivier Palombi.

Our long term goal is to develop high-level models helping users to express and convey their own narrative content (from fiction stories to more practical educational or demonstrative scenarios). Before being able to specify the narration, a first step is to define models able to express some a priori knowledge on the background scene and on the object(s) or character(s) of interest. Our first goal is to develop 3D ontologies able to express such knowledge. The second goal is to define a representation for narration, to be used in future storyboarding frameworks and virtual direction tools. Our last goal is to develop high-level models for virtual cinematography such as rule-based cameras able to automatically follow the ongoing action and semi-automatic editing tools enabling to easily convey the narration via a movie.

6.3.1. Virtual direction tools

Participants: Adela Barbulescu, Rémi Ronfard.



Figure 11. Live demo at the EXPERIMENTA exhibition in Grenoble.

During the third year of Adela Barbulescu's PhD thesis, we proposed a solution for converting a neutral speech animation of a virtual actor (talking head) to an expressive animation. Using a database of expressive audiovisual speech recordings, we learned generative models of audiovisual prosody for 16 dramatic attitudes (seductive, hesitant, jealous, scandalized, etc.) and proposed methods for transferring them to novel examples. Our results demonstrate that the parameters which describe an expressive performance present person-specific signatures and can be generated using spatio-temporal trajectories; parameters such as voice spectrum can be obtained at frame-level, while voice pitch, eyebrow raising or head movement depend both on the frame and the temporal position at phrase-level. This work was presented at the first joint conference on facial animation and audio-visual speech processing [16] and in a live demo at the EXPERIMENTA exhibition in Grenoble, and was seen by 1200 visitors.

6.3.2. Virtual cinematography

Participants: Quentin Galvane, Rémi Ronfard.

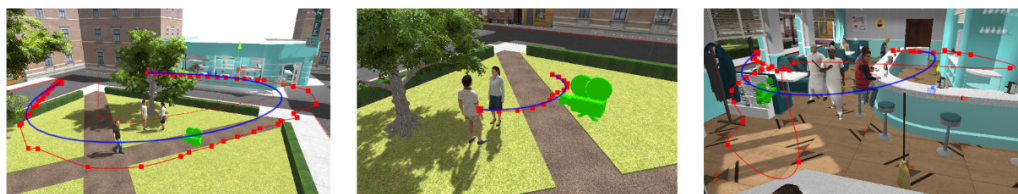


Figure 12. Automatic camera trajectory in [18].

During the third year of Quentin Galvane's PhD thesis, we proposed a solution for planning complex camera trajectories in crowded animation scenes [2] Galvane [18]. This work was done in a collaboration with Marc Christie in Rennes.

We also published new results from Vineet Gandhi's PhD thesis (defended in 2014) on the generation of cinematographic rushes from single-view recordings of theatre performances [32]. In that paper, we demonstrate how to use our algorithms to generate a large range of dynamic shot compositions from a single static view, a process which we call "vertical editing". Our patent application on this topic was reviewed positively and is being extended.

Those techniques were used to automatically generated cinematographically pleasant rushes from a monitor camera during rehearsals at Theatre des Celestins, as part of ANR project "Spectacle-en-Lignes". Results of the projects are described in two papers [28], Steiner [30] and we presented them to a professional audience during the Avignon theatre festival. This work was done in a collaboration with the Institut de Recherche et d'Innovation (IRI) at Centre Pompidou and the SYLEX team at LIRIS.

6.3.3. Film editing & narrative design

Participants: Quentin Galvane, Rémi Ronfard.

We proposed a new computational model for film editing at the AAAI artificial intelligence conference, which is based on semi-Makov chains [20]. Our model significantly extends previous work by explicitly taking into account the crucial aspect of timing (pacing) in film editing. Our proposal is illustrated with a reconstruction of a famous scene of the movie "Back to the future" in 3D animation, and a comparison of our automatic film editing algorithms with the director's version. Results are further discussed in two companion papers [31], Galvane [19]. This work was done in a collaboration with Marc Christie in Rennes. Future work is being planned to extend this important work to the case of live-action video (as described in the previous section) and to generalize for the case non-linear film editing including temporal ellipses and flashbacks.



Figure 13. Example of automatic movie sequence editing in [20].

6.4. Creating and Interacting with Virtual Prototypes

- **Scientist in charge:** Jean-Claude Léon.
- **Other permanent researchers:** Marie-Paule Cani, Olivier Palombi, Damien Rohmer, Rémi Ronfard.

The challenge is to develop more effective ways to put the user in the loop during content authoring. We generally rely on sketching techniques for quickly drafting new content, and on sculpting methods (in the sense of gesture-driven, continuous distortion) for further 3D content refinement and editing. The objective is to extend these expressive modeling techniques to general content, from complex shapes and assemblies to animated content. As a complement, we are exploring the use of various 2D or 3D input devices to ease interactive 3D content creation.

6.4.1. Sculpting shape hierarchies

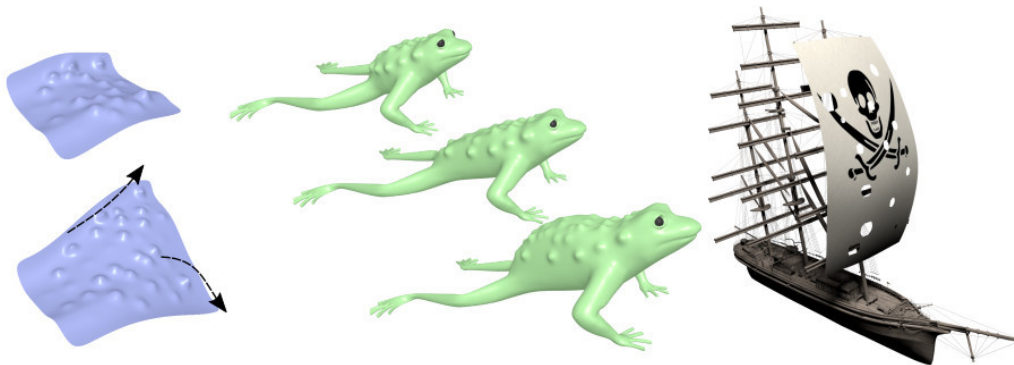


Figure 14. Our method in [13] allows surface stretching while the geometrical details are continuously duplicated.

Sculpting paradigm has been successfully applied to deform simple smooth surfaces. More complex objects representing virtual characters or real-life objects are however modeled as hierarchy of shapes with elements, sub-elements and details. Applying sculpting deformation to such objects is challenging as every parts of the hierarchy should stay coherent through the deformation.

When an object can be represented as a smooth underlying surface and a set of singular details, we proposed a real-time deformation approach enabling to freely stretch or squeeze the 3D object while continuously maintaining the details' appearance. Instead of stretching or squeezing the details the same way than the smooth underlying surface, we duplicate or merge them smoothly while ensuring that the distribution of details has the same characteristic than the original one. We published this work in Shape Modeling International [13].

In the case of more general object hierarchies, we are developing a new methodology to apply generic deformation into complex assemblies while preserving their properties in extending the shape grammar approach into our new *deformation grammar*. We presented our preliminary results as a communication in the GTMG conference [35].

6.4.2. Sketching and sculpting Virtual Worlds

Participants: Marie-Paule Cani, Guillaume Cordonnier, Ulysse Vimont.

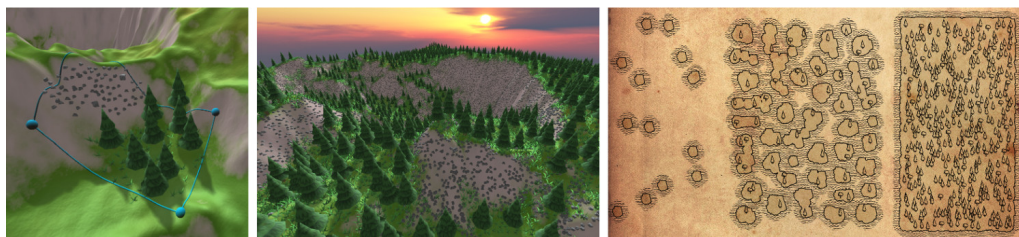


Figure 15. Example of maps generated using our World Brush method [8].

Modeling virtual worlds is particular challenging: the fractal-like distribution of details in terrain shapes makes them easy to identify, but very difficult to design using standard modeling software, even for expert users. Moreover, virtual worlds involve distributions of different categories of contents over terrains, such as vegetation, houses, roads or rivers. Efficiently modeling these sets of elements, which are statistically correlated, is indeed a challenge.

This year, our contributions to tackle these issues were two-folds:

Firstly, we investigated the use of a plate tectonics metaphor to generate plausible terrains from a simple vector map representing the location of the main rivers and mountain picks. The method uses a Voronoi tessellation of pick locations to automatically generate tectonic plates which themselves drive terrain folds. Hydraulic erosion is then used to further sculpt the terrain and add details, while the specified rivers are considered to maintain consistency with the input map. This work was published in [27]. A more accurate modeling of large scale fluvial erosion and plates tectonics phenomena was investigated in Guillaume Cordonnier's master thesis and is the object of his PhD, which started in October 2015.

Secondly, we proposed a paint-based interface to tackle the problem of easily populating a terrain with distributions of objects (trees, rocks, grass, houses, etc) or of graph-like structures such as rivers and roads. The key point of our solution is to learn statistics about distributions of elements and their correlation with other distributions, with graph structures, or with terrain slope, and store the resulting histograms as "colors" in a palette interface. After creating a few local distribution manually, the user selects them with a pipette tool, and is able to reuse them with a brush. We also provided a gradient tool to interpolate between two such "colors" and a move tool enabling, for instance to move groups of trees and rocks over a terrain while maintaining the adequate correlation with local slope, and a deformation interface based on seam carving enabling to seamlessly stretch or compress a region of virtual world. This work, a collaboration between Arnaud Emilien when defended his PhD in December 2014, Ulysse Vimont, Marie-Paule Cani, and Bedrich Benes from Purdue University, was published at Siggraph 2015 [8].

6.4.3. Sketching and sculpting Motion

Participants: Marie-Paule Cani, Martin Guay, Kevin Jordao, Rémi Ronfard.

Sketching and sculpting methods were restricted so far to the design of static shapes. One of our research goals has been to extend these interaction metaphors to motion design. This year, this included three specific contributions.

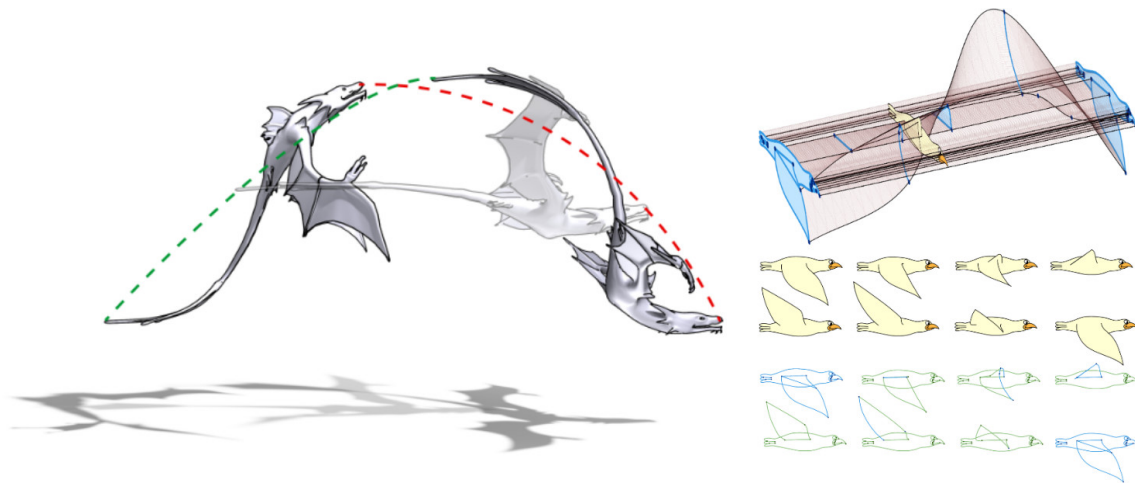


Figure 16. Left: Result of our Sculpting Motion method [9]. Right: Result of our Vector Graphics method enabling to design vectorial animation [6].

Firstly, to handle sketch-based representation of motion in the 2D case, we extended the static vector graphics complex data structure, which we had introduced at Siggraph last year, to vector graphics animations with time-varying topology [6]. This second paper was presented at Siggraph again this year. The proposed data structure is augmented with a rich set of editing operations, which can be used to quickly interpolate 2D drawings with different topologies. This work was done within a collaboration with Boris Dalstein and Michiel van de Panne from UBC, Canada.

Secondly, following a first method enabling to sculpt crowd animations (Jordao, Eurographics 2014), we developed a painting interface enabling to specify both density and main directions of motion in an animated crowd. The resulting system is still based on crowd-patches, i.e. the crowd motion is an assembly of local trajectories defined in interconnected patches. Our new painting system, called Crowd-Art, uses discrete changes in loop trajectories to evolve the number of in/out constraints in each patch until the requested density and directions are best matched. See [22]. This concluded Kevin Jordao's PhD thesis, co-advised by Julien Pettre from the MimeTIC team and in collaboration with Marc Christie, defended in December 2015.

Lastly, we developed the first expressive interface to interactively sketch and progressively sculpt and refine character motion. Our solution is based on a space-time sketching metaphor: The user sketches a single space-time stroke, which is used to initialize a series of dynamic lines of action, serving as intermediates to animate the character's model. Motion and shape deformation can be immediately replayed from this single stroke, since it sets at the same time shape, trajectory and speed (defined from the drawing speed). Thanks to visual feedback, the user can easily refine the resulting motion by editing specific lines of actions at fixed times, or by composing several motions together. This work, published at Siggraph, is one of the first methods enabling arbitrary motion to be defined from scratch by a beginner [9]. Together to another work enabling to add dynamics to character motion [21], this concluded Martin Guay's PhD thesis, defended in June 2015.

LAGADIC Project-Team

7. New Results

7.1. Visual tracking

7.1.1. *Object detection*

Participant: Eric Marchand.

We addressed the challenge of detecting and localizing a poorly textured known object, by initially estimating its complete 3D pose in a video sequence [45]. Our solution relies on the 3D model of the object and synthetic views. The full pose estimation process is then based on foreground/background segmentation and on an efficient probabilistic edge-based matching and alignment procedure with the set of synthetic views, classified through an unsupervised learning phase. Our study focuses on space robotics applications and the method has been tested on both synthetic and real images, showing its efficiency and convenience, with reasonable computational costs.

7.1.2. *Registration of multimodal images*

Participant: Eric Marchand.

This study has been realized in collaboration with Brahim Tamadazte and Nicolas Andreff from Femto-ST, Besançon. Following our developments in visual tracking and visual servoing from the mutual information [3], it concerned mutual information-based registration of white light images vs. fluorescence images for micro-robotic laser microphonosurgery of the vocal folds. Nelder-Mead Simplex for nonlinear optimization has been used to minimize the cost-function [43].

7.1.3. *Pose estimation from RGB-D sensor*

Participant: Eric Marchand.

RGB-D sensors have become in recent years a product of easy access to general users. They provide both a color image and a depth image of the scene and, besides being used for object modeling, they can also offer important cues for object detection and tracking in real-time. In this context, the work presented in this paper investigates the use of consumer RGB-D sensors for object detection and pose estimation from natural features. Two methods based on depth-assisted rectification are proposed, which transform features extracted from the color image to a canonical view using depth data in order to obtain a representation invariant to rotation, scale and perspective distortions. While one method is suitable for textured objects, either planar or non-planar, the other method focuses on texture-less planar objects [18]

7.1.4. *3D localization for airplane landing*

Participants: Noël Mériaux, François Chaumette, Patrick Rives, Eric Marchand.

This study is realized in the scope of the ANR VisioLand project (see Section 9.2.2). In a first step, we have considered and adapted our model-based tracker [2] to localize the aircraft with respect to the airport surroundings. Satisfactory results have been obtained from real image sequences provided by Airbus. In a second step, we have started to perform this localization from a set of keyframe images corresponding to the landing trajectory.

7.2. Visual servoing

7.2.1. *Histogram-based visual servoing*

Participants: Quentin Bateux, Eric Marchand.

Classically visual servoing considers the regulation in the image of a set of visual features (usually geometric features). Direct visual servoing schemes, such as photometric visual servoing, have been introduced in order to consider every pixel of the image as a primary source of information and thus avoid the extraction and the tracking of such geometric features. This year, we proposed a method to extend these works by using a global descriptor, namely intensity histograms, on the whole or multiple sub-sets of the images in order to achieve control of a 6 degrees of freedom (DoF) robot [30][53].

7.2.2. *Photometric moment-based visual servoing*

Participants: Manikandan Bakthavatchalam, François Chaumette.

This work also belongs to the class of direct visual servoing. Its goal was to use photometric moments as visual features in order to increase the convergence domain of this approach by reducing the non linearity of the control problem. In order to cope with appearance and disappearance of some parts of the environment during the camera motion, a spatial weight has been introduced in the definition of photometric moments. Thanks to a particular design of this weight, the analytical form of the interaction matrix has been obtained, from which it was possible to select a set of moment combinations to control all the six degrees of freedom of the system. Satisfactory experimental results have been obtained [29][8], even if the loss of invariance properties makes the optimal design of visual features still an open problem.

7.2.3. *Model predictive visual servoing*

Participants: Nicolas Cazy, Paolo Robuffo Giordano, François Chaumette.

The goal of this work is to exploit Model Predictive Control (MPC) techniques for dealing in a robust way with loss of features during a IBVS task. The work [31] provides an experimental validation of different correction schemes able to cope with loss of features due to occlusions of limited camera field of view. The reported results show the effectiveness of the proposed techniques during the servoing of four point features.

7.2.4. *Nanomanipulation*

Participants: Le Cui, Eric Marchand.

Following our work related to scanning electron microscope (SEM) calibration [12] we considered the control of a micro robot using a direct photometric visual servoing that uses only the pure image information as a visual feature, instead of using classic geometric features such as points or lines. However, in micro-scale, using only image intensity as a visual feature performs unsatisfactorily in cases where the photometric variation is low, such as motions along vision sensor's focal axis under a high magnification. In order to improve the performance and accuracy in those cases, an approach using hybrid visual features is proposed in this paper. Image gradient is employed as a visual feature on z axis while image intensity is used on the other 5 DoFs to control the motion. A 6-DoF micro-positioning task is accomplished by this hybrid visual servoing scheme [34].

We also considered a full scale autofocus approach for SEM [35]. The optimal focus (in-focus) position of the microscope is achieved by maximizing the image sharpness using a vision-based closed-loop control scheme. An iterative optimization algorithm has been designed using the sharpness score derived from image gradient information. The proposed method has been implemented and validated using a tungsten gun SEM at various experimental conditions like varying raster scan speed, magnification at real-time.

7.2.5. *Audio-based control*

Participants: Aly Magassouba, François Chaumette.

This study is not concerned with visual servoing, but to the application of the same principle of sensor-based control to audio sensors. It is made in collaboration with Nancy Bertin from Panama group at IriSa, Inria Rennes-Bretagne Atlantique. In a first step, we have determined the analytical form of the interaction matrix of audio features based on the time difference of arrival on two microphones. From this modeling step, we have determined the different virtual linkages that can be realized in function of the number and configuration of sources [41]. First experimental results using two microphones mounted on the Pioneer mobile robot (see Section 6.9) have been recently obtained.

7.3. Visual navigation of mobile robots

7.3.1. Visual navigation from straight lines

Participants: Suman Raj Bista, Paolo Robuffo Giordano, François Chaumette.

This study is concerned with visual autonomous navigation in indoor environments. As in our previous works concerning navigation outdoors [4], the approach is based on a topological localization of the current image with respect to a set of keyframe images, but the visual features used for this localisation as well as for the visual servoing is not based on points of interest, but straight lines that are more common indoors. Satisfactory experimental results have been obtained using the Pioneer mobile robot (see Section 6.9) [23].

7.3.2. Autonomous navigation of a wheelchair and social navigation

Participants: Vishnu Karakkat Narayanan, François Pasteau, Marie Babel.

Navigating within an unknown indoor environment using an electric wheelchair is a challenging task, especially if the user suffers from severe disabilities. We presented in [22] a framework for vision-based autonomous indoor navigation in an electric wheelchair capable of following corridors, and passing through open doorways using a single doorpost. The designed control schemes have been implemented onto a robotized wheelchair and experimental results show the robust behaviour of the designed system.

We then introduced in [40] a task-based control law which can serve as a low-level system for equitably joining interacting groups, while conforming to social conventions. The system uses the position and orientation of the participating humans with respect to a rigid sensor frame in order to control the translational and rotational velocity of a wheelchair so that the robot positions itself aptly at the meeting point

7.3.3. Semi-autonomous control of a wheelchair for navigation assistance

Participants: Vishnu Karakkat Narayanan, François Pasteau, Marie Babel.

To address the wheelchair driving assistance issue, we proposed in [56][28] a unified shared control framework able to smoothly correct the trajectory of the electrical wheelchair. The system integrates the manual control with sensor-based constraints by means of a dedicated optimization strategy. The resulting low-complex and low-cost embedded system is easily plugged onto on-the-shelf wheelchairs.

The robotic solution has been then validated through clinical trials that have been conducted within the Rehabilitation Center of Pôle Saint Hélier (France) with 25 volunteering patients presenting different disabling neuro-pathologies. This assistive tool is shown to be intuitive and robust as it respects the user intention, it does not alter perception while reducing the number of collisions in case of hazardous maneuvers or in crowded environment [27].

7.4. 3D Scene Mapping

7.4.1. Structure from motion

Participants: Riccardo Spica, Paolo Robuffo Giordano, François Chaumette.

Structure from motion (SfM) is a classical and well-studied problem in computer and robot vision, and many solutions have been proposed to treat it as a recursive filtering/estimation task. However, the issue of *actively* optimizing the transient response of the SfM estimation error has not received a comparable attention. In the work [50] we have addressed the active estimation of the 3D structure of an observed planar scene by comparing three different techniques: a homography decomposition (a well-established method taken as a baseline), a least-square fitting of a reconstructed 3D point cloud, and a direct estimation based on the observation of a set of discrete image moments made of a collection of image points belonging to the observed plane. The experimental results confirmed the importance of actively controlling the camera motion in order to obtain a faster convergence for the estimation error, as well as the superiority of the third method based on the machinery of image moments for what concerns robustness against noise and outliers. In [51] the active estimation scheme has been improved by considering a set of features invariant to camera rotations. This

way, the dynamics of the structure estimation becomes independent of the camera angular velocity whose measurement is, thus, no longer required for implementing the active SfM scheme. Finally, in [46] the issue of determining online the ‘best’ combination of image moments for reconstructing the scene structure has been considered. By defining a new set of weighted moments as a weighted sum of traditional image moments, it is indeed possible to optimize for the weights online during the camera motion. The SfM scheme then automatically selects online the best combination of image moments to be used as measurements as a function of the current scene.

7.4.2. Scene Registration based on Planar Patches

Participants: Eduardo Fernandez Moral, Patrick Rives.

Scene registration consists of estimating the relative pose of a camera with respect to a scene previously observed. This problem is ubiquitous in robot localization and navigation. We propose a probabilistic framework to improve the accuracy and efficiency of a previous solution for structure registration based on planar representation. Our solution consists of matching graphs where the nodes represent planar patches and the edges describe geometric relationships. The maximum likelihood estimation of the registration is estimated by computing the graph similarity from a series of geometric properties (areas, angles, proximity, etc..) to maximize the global consistency of the graph. Our technique has been validated on different RGB-D sequences, both perspective and spherical [14].

7.4.3. Robust RGB-D Image Registration

Participants: Tawsif Gokhool, Renato José Martins, Patrick Rives.

Estimating dense 3D maps from stereo sequences remains a challenging task where building compact and accurate scene models is relevant for a number of tasks, from localization and mapping to scene rendering [20], [10]. In this context, this work deals with generating complete geometric and photometric “minimal” model of indoor/outdoor large-scale scenes, which are stored within a sparse set of spherical images to asset photo-geometric consistence of the scene from multiple points-of-views. To this end, a probabilistic data association framework for outlier rejection is formulated, enhanced with the notion of landmark stability over time. The approach was evaluated within the frameworks of image registration, localization and mapping, demonstrating higher accuracy and larger convergence domains over different datasets [39].

7.4.4. Accurate RGB-D Keyframe Representation of 3D Maps

Participants: Renato José Martins, Eduardo Fernandez Moral, Patrick Rives.

Keyframe-based maps are a standard solution to produce a compact map representation from a continuous sequence of images, with applications in robot localization, 3D reconstruction and place recognition. We have present a approach to improve keyframe-based maps of RGB-D images based on two main filtering stages: a regularization phase in which each depth image is corrected considering both geometric and photometric image constraints (planar and superpixel segmentation); and a fusion stage in which the information of nearby frames (temporal continuity of the sequence) is merged (using a probabilistic framework) to improve the accuracy and reduce the uncertainty of the resulting keyframes. As a result, more compact maps (with less keyframes) are created. We have validated our approach with different kind of RGB-D data including both indoor and outdoor sequences, and spherical and perspective sensors, demonstrating that our approach compares and outperforms the state-of-the-art [42].

7.4.5. Semantic Representation For Navigation In Large-Scale Environments

Participants: Romain Drouilly, Patrick Rives.

Autonomous navigation is one of the most challenging problem to address to allow robots to evolve in our everyday environments. Map-based navigation has been studied for a long time and researches have produced a great variety of approaches to model the world. However, semantic information has only recently been taken into account in those models to improve robot efficiency.

Mimicking human navigation is a challenging goal for autonomous robots. This requires to explicitly take into account not only geometric representation but also high-level interpretation of the environment [9]. We propose a novel approach demonstrating the capability to infer a route in a global map by using semantics. Our approach relies on an object-based representation of the world automatically built by robots from spherical images. In addition, we propose a new approach to specify paths in terms of high-level robot actions. This path description provides robots with the ability to interact with humans in an intuitive way. We perform experiments on simulated and real-world data, demonstrating the ability of our approach to deal with complex large-scale outdoor environments whilst dealing with labelling errors [37].

Mapping evolving environments requires an update mechanism to efficiently deal with dynamic objects. In this context, we propose a new approach to update maps pertaining to large-scale dynamic environments with semantics. While previous works mainly rely on large amount of observations, the proposed framework is able to build a stable representation with only two observations of the environment. To do this, scene understanding is used to detect dynamic objects and to recover the labels of the occluded parts of the scene through an inference process which takes into account both spatial context and a class occlusion model. Our method was evaluated on a database acquired at two different times with an interval of three years in a large dynamic outdoor environment. The results point out the ability to retrieve the hidden classes with a precision score of 0.98. The performances in term of localisation are also improved [36].

7.5. Control of single and multiple Unmanned Aerial Vehicles

7.5.1. Single UAV

Participant: Paolo Robuffo Giordano.

Over the last years the robotics community witnessed an increasing interest in the Unmanned Aerial Vehicle (UAV) field. In particular quadrotor UAVs have become more and more widespread in the community as experimental platform for, e.g., testing novel 3D planning, control and estimation schemes in real-world indoor and outdoor conditions. Indeed, in addition to being able to take-off and land vertically, quadrotors can reach high angular accelerations thanks to the relatively long lever arm between opposing motors. This makes them more agile than most standard helicopters or similar rotorcraft UAVs, and thus very suitable to realize complex tasks such as aerial mapping, air pollution monitoring, traffic management, inspection of damaged buildings and dangerous sites, as well as agricultural applications such as pesticide spraying.

Despite these clear advantages, a clear shortcoming of the quadrotor design lies in its inherent underactuation (only 4 actuated propellers for the 6 dofs of the quadrotor pose). This underactuation limits the quadrotor flying ability in free or cluttered space and, furthermore, it also degrades the possibility of interacting with the environment by exerting desired forces in arbitrary directions. In [24], a novel design for a quadrotor UAV with tilting propellers which is able to overcome these limitations has been presented and experimentally validated. Indeed, the additional set of 4 control inputs actuating the propeller tilting angles can be shown to yield full actuation to the quadrotor position/orientation in space, thus allowing it to behave as a fully-actuated flying vehicle and to overcome the aforementioned underactuation problem.

Furthermore, the issue of estimating online the UAV self-motion from vision has been considered. To this end, a novel nonlinear estimation scheme able to recover the metric UAV linear velocity from the *scaled* one obtained from the decomposition of the optical flow has been proposed in [15]. The observability conditions (in terms of persistency of excitation) needed to ensure a converging estimation have also been studied. The reported experimental results confirmed the effectiveness of the estimation scheme in recovering a reliable and accurate estimation of the UAV self-motion (linear and angular velocities) in realistic conditions.

This work has been realized in collaboration with the Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

7.5.2. Collective control of multiple UAVs

Participants: Fabrizio Schiano, Paolo Robuffo Giordano.

The challenge of coordinating the actions of multiple robots is inspired by the idea that proper coordination of many simple robots can lead to the fulfilment of arbitrarily complex tasks in a robust (to single robot failures) and highly flexible way. Teams of multi-robots can take advantage of their number to perform, for example, complex manipulation and assembly tasks, or to obtain rich spatial awareness by suitably distributing themselves in the environment. Within the scope of robotics, autonomous search and rescue, firefighting, exploration and intervention in dangerous or inaccessible areas are the most promising applications.

In the context of multi-robot (and multi-UAV) coordinated control, *connectivity* of the underlying graph is perhaps the most fundamental requirement in order to allow a group of robots accomplishing common goals by means of *decentralized* solutions. In fact, graph connectivity ensures the needed continuity in the data flow among all the robots in the group which, over time, makes it possible to share and distribute the needed information. However, connectivity alone is not sufficient to perform certain tasks when only *relative sensing* is used. For these systems, the concept of *rigidity* provides the correct framework for defining an appropriate sensing and communication topology architecture. Rigidity is a combinatorial theory for characterizing the “stiffness” or “flexibility” of structures formed by rigid bodies connected by flexible linkages or hinges. In a broader context, rigidity turns out to be an important architectural property of many multi-agent systems when a common inertial reference frame is unavailable. Applications that rely on sensor fusion for localization, exploration, mapping and cooperative tracking of a target, all can benefit from notions in rigidity theory. The concept of rigidity, therefore, provides the theoretical foundation for approaching decentralized solutions to the aforementioned problems using distance measurement sensors, and thus establishing an appropriate framework for relating system level architectural requirements to the sensing and communication capabilities of the system.

In [26], a decentralized gradient-based rigidity maintenance action for a group of quadrotor UAVs has been proposed and tested in real experimental conditions. By starting in a rigid configuration, the group of UAVs is able to estimate their relative position from sole relative distance measurements, and then use these estimated relative positions in a control action able to preserve rigidity of the whole formation despite presence of sensor limitations (maximum range and line-of-sight occlusions), possible collisions with obstacles and inter-robot collisions. Furthermore, in [52] the novel case of *bearing rigidity* for directed graphs has been considered: here, rather than distances the measurements are the 3D bearing vectors expressed in the local body-frame of each agent. The theory has been developed for the case of planar agents in $SE(2)$ and a ‘scale-free’ bearing controller has been proposed, able to steer the robot group towards a desired bearing formation.

These works were realized in collaboration with the robotics group at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany and with Technion, Israel.

7.5.3. Cooperative localization using interval analysis

Participants: Vincent Drevelle, Ide Flore Kenmogne Fokam.

In the context of multi-robot fleets, cooperative localization consists in gaining better position estimate through measurements and data exchange with neighboring robots. Positioning integrity (i.e., providing reliable position uncertainty information) is also a key point for mission-critical tasks, like collision avoidance. The goal of this work is to compute position uncertainty volumes for each robot of the fleet, using a decentralized method (i.e using only local communication with the neighbors). The problem is addressed in a bounded-error framework, with interval analysis and constraint propagation methods. These methods enable to provide guaranteed position error bounds, assuming bounded-error measurements. They are not affected by over-convergence due to data incest, which makes them a well sound framework for decentralized estimation. Encouraging results have already been obtained for multi-robot underwater positioning with acoustical range measurements. Ongoing work focuses on cooperative localization in a multi-UAV fleet with image-based measurements (bearings).

7.6. Medical robotics

7.6.1. Non-rigid target tracking in ultrasound images combining dense information and physically-based model

Participants: Lucas Royer, Alexandre Krupa.

This study concerns the real-time tracking of deformable targets within a sequence of ultrasound (US) images. The proposed approach combines dense information with a physically-based model and has therefore the advantage of not using any fiducial marker. The physical model is represented by a mass-spring damper system driven by external and internal forces. The external forces are obtained by maximizing an image similarity metric between a reference target and the deformed target along the time. The internal forces of the mass-spring damper system constrain the deformation to be physically plausible and therefore efficiently reduce the sensitivity to the speckle noise. This approach was first validated from simulated and real sequences of 2D US images [49]. It was then extended for deformable target tracking in a sequence of 3D ultrasound volumes and tested on a robotic setup used to apply deformation on an organic phantom [48]. The performance of this deformable 3D target tracking approach was evaluated with visual assessment combined with robotic odometry ground truth. This method was also tested and compared with respect to state-of-the-art techniques by using 3D image databases provided by MICCAI CLUST'14 and CLUST'15 challenges [47] (MICCAI Challenge on Liver Ultrasound Tracking). It was awarded by the organizers of the CLUST challenges as being the best method for accurate target tracking in 3D ultrasound sequences. We recently improved our approach in order to increase its robustness to the presence of ultrasound shadows, local illumination changes and image occlusions.

7.6.2. 3D steering of flexible needle by ultrasound visual servoing

Participants: Pierre Chatelain, Jason Chevrier, Marie Babel, Alexandre Krupa.

The objective of this work is to provide robotic assistance during needle insertion procedures such as biopsy or ablation of localized tumor. In previous work, we designed a control approach based on a duty cycling technique for steering a beveled-tip flexible needle actuated by a robotic arm in such a way to control the needle curvature in 3D space and reach a desired target by visual servoing. In this preliminary work, the control approach was validated by using visual features extracted from 2 images provided by 2 orthogonal cameras observing a translucent gelatin phantom where the needle was inserted. This year, we have pursued our work towards this needle steering robotic assistance by developing a new algorithm able to track in real-time a flexible needle in a sequence of 3D ultrasound images (volumes). The flexible needle modeled as a polynomial curve is tracked during the automatic insertion using particle filtering. This new tracking algorithm enables real-time closed-loop needle control with 3D ultrasound feedback. The target to reach was manually defined by the user in the US image and can be on-line tracked thanks to the template tracking algorithm proposed in [21] based on ultrasound dense visual servoing [7]. Experimental results of an automatic needle tip positioning in a home-made gelatine phantom demonstrate the feasibility of 3D ultrasound-guided needle steering for reaching a desired target by ultrasound visual servoing [33]. Recently a new control law for needle steering that uses both direct manipulation of the needle base and the duty cycling method has been studied. It is based on a 3D model of a beveled tip needle using virtual springs that characterize the needle mechanical interaction with soft tissue. From this model, a measure of the controllability of the needle tip degrees of freedom was proposed in order to mix the control between the direct base manipulation and the duty cycling technique. Preliminary simulations show that this hybrid control allows better targeting capabilities in terms of larger needle workspace and reduced needle bending.

7.6.3. Optimization of ultrasound image quality by visual servoing

Participants: Pierre Chatelain, Alexandre Krupa.

This study focuses on a new ultrasound-based visual servoing approach that optimizes the positioning of an ultrasound probe manipulated by a robotic arm in order to improve the quality of the acquired ultrasound images. To this end, we use the recent framework of ultrasound confidence map, developed in the Chair for Computer Aided Medical Procedures and Augmented Reality of Prof. Nassir Navab, which aims at estimating the per-pixel quality of the ultrasound signal based on a model of sound propagation in soft tissues. More specifically, we treat the ultrasound confidence maps as a new modality and designed a visual servoing control law for image quality optimization. We illustrated our approach with the application of robotic tele-echography where the in-plane rotation of a 2D probe is visually servoed by the confidence map and the other degrees of

freedom are teleoperated by the user. Experiments performed on both an ultrasound examination training phantom and ex vivo tissue samples validated this new concept [32]. Currently, we consider the confidence-driven servoing of other degrees of freedom, in particular out-of-plane motions that were controlled in our previous works from image moments [6], which could provide finer control of the image quality.

7.6.4. Visual servoing based on ultrasound elastography

Participants: Pedro Alfonso Patlan Rosales, Alexandre Krupa.

This study concerns the use of the ultrasound elastography as a new image modality for the control of the motion of an ultrasound probe actuated by a robotic manipulator. Elastography imaging is performed by applying continuous stress variation on soft tissues in order to estimate a strain map of the observed tissues. It is obtained by estimating, from the RF (radio-frequency) signal along each scan line of the probe transducer, the echo time delays between pre- and post-compressed tissue. Usually, this continuous stress variation is performed manually by the user who manipulates the US probe and it results therefore in a user-dependent quality of the elastography image. To improve the US elastography imaging, we recently developed an assistant robotic palpation system that automatically moves an ultrasound probe in such a way to optimize ultrasound elastography. The main originality of this preliminary work concerns the use of the elastography modality directly as input of the robot controller thanks to an innovative ultrasound elastography-based visual servoing approach.

7.6.5. Visual servoing using shearlet transform

Participants: Lesley-Ann Dufлот, Alexandre Krupa.

Similar to wavelet transform, shearlet transform is usually used in the field of signal or image compression. At the best of our knowledge these image representations were never used directly as feedback of a closed-loop control scheme. The objective of this work is to study the feasibility of using the coefficients of shearlet transform of the observed ultrasound image directly as the visual features of an image-based visual servoing. In this study we estimated numerically the interaction matrix that links the time variation of the coarsest coefficients of the shearlet to the motion of the ultrasound probe. This shearlet-based visual servoing was experimentally tested for automatically positioning a 2D US probe, held by a robot, on a desired section of an abdominal phantom. The first results demonstrated promising performances.

LARSEN Team

7. New Results

7.1. Lifelong Autonomy

7.1.1. Adaptation / Learning

Participant: Jean-Baptiste Mouret.

We collaborate on this subject with Jeff Clune (University of Wyoming, USA).

7.1.1.1. Adaptation to Unforeseen Damage Conditions

Whereas animals can quickly adapt to injuries, current robots cannot “think outside the box” to find a compensatory behaviour when they are damaged: they are limited to their pre-specified self-sensing abilities and can diagnose only anticipated failure modes, an impracticality for complex robots. A promising approach to reducing robot fragility involves having robots learn appropriate behaviours in response to damage, but current techniques are slow even with small, constrained search spaces. We introduced an intelligent trial-and-error algorithm that allows robots to adapt to damage in less than two minutes in large search spaces without requiring self-diagnosis or pre-specified contingency plans [11]. Before the robot is deployed, it uses a novel technique (based on evolutionary algorithms) to create a detailed map of the space of high-performing behaviours. This map represents the robot’s prior knowledge about what behaviours it can perform and their value. When the robot is damaged, it uses this prior knowledge to guide a trial-and-error learning algorithm (based on Bayesian optimization) that conducts intelligent experiments to rapidly discover a behaviour that compensates for the damage. Experiments reveal successful adaptations for a legged robot injured in five different ways, including damaged, broken, and missing legs, and for a robotic arm with joints broken in 14 different ways. This new algorithm will enable more robust, effective, autonomous robots, and may shed light on the principles that animals use to adapt to injury.

This work was the cover of Nature on the 28th of May, 2015 (see the “highlights” section).

7.1.2. Robotics Perception

Participants: François Charpillat, Francis Colas, Abdallah Dib, Van Quan Nguyen.

We collaborate on this subject with Emmanuel Vincent from the Multispeech team (Inria Nancy - Grand Est).

7.1.2.1. Audio Source Localization

We considered, here, the task of audio source localization using a microphone array on a mobile robot. Active localization algorithms have been proposed in the literature that can estimate the 3D position of a source by fusing the measurements taken for different poses of the robot. However, the robot movements are typically fixed or they obey heuristic strategies, such as turning the head and moving towards the source, which may be suboptimal. This work proposes an approach to control the robot movements so as to locate the source as quickly as possible [17]. We represent the belief about the source position by a discrete grid and we introduce a dynamic programming algorithm to find the optimal robot motion minimizing the entropy of the grid. We report initial results in a real environment.

This work is carried on through the PhD Thesis of Van Quan Nguyen under the supervision of Emmanuel Vincent and Francis Colas.

7.1.2.2. State Estimation for Autonomous Surface Vessels

Autonomous Surface Vessels (ASVs) are increasingly proposed as tools to automatize environmental data collection, bathymetric mapping and shoreline monitoring. For many applications it can be assumed that the boat operates on a 2D plane. However, with the involvement of exteroceptive sensors like cameras or laser rangefinders, knowing the 3D pose of the boat becomes critical. We formulated three different algorithms based on 3D extended Kalman filter (EKF) state estimation for ASVs localization [12]. We compared them using field testing results with ground truth measurements, and demonstrated that the best performance is achieved with a model-based solution in combination with a complementary filter for attitude estimation. Furthermore, we presented a parameter identification methodology and showed that it also yielded accurate results when used with inexpensive sensors. Finally, we presented a long-term series (i.e., over a full year) of shoreline monitoring data sets and discussed the need for map maintenance routines based on a variant of the Iterative Closest Point (ICP) algorithm.

7.1.2.3. Geometric Registration

We proposed a review of geometric registration in robotics [16]. Registration algorithms associate sets of data into a common coordinate system. They have been used extensively in object reconstruction, inspection, medical application, and localization of mobile robotics. We focus on mobile robotics applications in which point clouds are to be registered. While the underlying principle of those algorithms is simple, many variations have been proposed for many different applications. In this work, we gave a historical perspective of the registration problem and showed that the plethora of solutions can be organized and differentiated according to a few elements. Accordingly, we presented a formalization of geometric registration and cast algorithms proposed in the literature into this framework. Finally, we reviewed a few applications of this framework in mobile robotics that cover different kinds of platforms, environments, and tasks. These examples allowed us to study the specific requirements of each use case and the necessary configuration choices leading to the registration implementation. Ultimately, the objective of this work is to provide guidelines for the choice of geometric registration configuration.

7.1.2.4. Robust Dense Visual Odometry for RGB-D Cameras in a Dynamic Environment

Visual odometry is a fundamental challenge in robotics and computer vision. The aim of our work is to estimate RGB-D camera motion (onboard a mobile robot) from RGB-D images in a dynamic scene with people moving in the scene. Most of the existing methods have a poor localization performance in such case, which makes them inapplicable in real world conditions. This year, we have proposed a new dense visual odometry method [27] that uses random sampling consensus (RANSAC) to cope with dynamic scenes. We show the efficiency and robustness of the proposed method on a large set of experiments in challenging situations and from publicly available benchmark datasets. Additionally, we compare our approach to another state-of-art method based on M-estimator that is used to deal with dynamic scenes. Our method gives similar results on benchmark sequences and better results on our own dataset.

7.1.3. Distributed Sensing and Acting

Participants: Mihai Andries, Amine Boumazza, François Charpillet, Iñaki Fernández Pérez, Nassim Kaldé.

We collaborate on this subject with Olivier Simonin from the Chroma team (Inria Grenoble - Rhône Alpes).

7.1.3.1. Localisation of Humans, Objects and Robots Interacting on Load-Sensing Floors

The use of floor sensors in ambient intelligence contexts began in the late 1990's. We designed such a sensing floor in Nancy in collaboration with Hikob company (<http://www.hikob.com/>) and Inria SED (*service d'expérimentation et de développement*). This is a load-sensing floor which is composed of square tiles, each equipped with two ARM processors (Cortex m3 and a8), 4 load cells, and a wired connection to the four neighboring cells. Ninety tiles cover the floor of our intelligent apartment experimental platform. This load-sensing floor includes as well a LED lighting system which sits flush with the floor surface. This provides people with a new way to interact with their environment at home. This year, we have focused on localisation, tracking and recognition of humans, objects and robots interacting on load-sensing floors [9]. Inspired by computer vision, the proposed technique processes the floor pressure-image by segmenting

the blobs containing objects, tracking them, and recognizing their contents through a mix of inference and combinatorial search. The result lists the probabilities of assignments of known objects to observed blobs. The concept was successfully evaluated in daily life activity scenarii, involving multi-object tracking and recognition on low resolution sensors, crossing of user trajectories, and weight ambiguity.

7.1.3.2. *Online Distributed Learning for a Swarm of Robots*

We propose a novel innovation marking method [22] for neuro-evolution of augmenting topologies in embodied evolutionary robotics. This method does not rely on a centralized clock, which makes it well suited for the decentralized nature of embodied evolution where no central evolutionary process governs the adaptation of a team of robots exchanging messages locally. This method is inspired from event dating algorithms, based on logical clocks, that are used in distributed systems, where clock synchronization is not possible. We compare our method to odNEAT, an algorithm in which agents use local time clocks as innovation numbers, on two multi-robot learning tasks: navigation and item collection. Our experiments showed that the proposed method performs as well as odNEAT, with the added benefit that it does not rely on synchronization of clocks and is not affected by time drifts.

The effect of selection pressure on evolution in centralized evolutionary algorithms (EA's) is relatively well understood. Selection pressure pushes evolution toward better performing individuals. However, distributed EA's in an Evolutionary Robotics (ER) context differ in that the population is distributed across the agents, and a global vision of all the individuals is not available. In this work, we analyze the influence of selection pressure in such a distributed context. We propose a version of mEDEA [22] that adds a selection pressure, and evaluate its effect on two multi-robot tasks: navigation and obstacle avoidance, and collective foraging. Experiments show that even small intensities of selection pressure lead to good performances, and that performance increases with selection pressure. This is opposed to the lower selection pressure that is usually preferred in centralized approaches to avoid stagnating in local optima.

7.1.3.3. *Online Distributed Exploration of an Unknown Environment by a Swarm of Robots*

This year, we have proposed a new taboo-list approach [18] for multi-robot exploration of unknown structured environments, in which robots are implicitly guided in their navigation on a globally shared map. Robots have a local view of their environment, inside which they navigate in an asynchronous manner. When the exploration is complete, robots gather at a rendezvous point. The novelty consists in using a distributed exploration algorithm which is not guided by frontiers to perform this task. Using the Brick and Mortar Improved ant-algorithm as a base, we add robot-perspective vision, variable vision range, and an optimization which prevents agents from going to the rendezvous point before exploration is complete. The algorithm was evaluated in simulation on a set of standard maps.

Another work [14] carried out within the PhD of Nassim Kaldé concerns exploration in populated environments. The difficulty here is that pedestrian flows can severely impact performances. However, humans have adaptive skills for taking advantage of these flows while moving. Therefore, in order to exploit these human abilities, we propose a novel exploration strategy that explicitly allows for human-robot interactions. Our model for exploration in populated environments combines the classical frontier-based strategy with our interactive approach. We implement interactions where robots can locally choose a human guide to follow and define a parametric heuristic to balance interaction and frontier assignments. Finally, we evaluate to which extent human presence impacts our exploration model in terms of coverage ratio, travelled distance and elapsed time to completion.

7.2. Natural Interaction with Robotics Systems

7.2.1. *Human Characterization*

Participants: François Charpillat, Abdallah Dib, Xuan Son Nguyen, Vincent Thomas.

We collaborate on this subject with Olivier Buffet and Alain Dutech from Inria Nancy - Grand Est, Arsène Fansi Tchango and Fabien Flacher from Thales ThereSIS, and Alain Filbois from SED Inria Nancy - Grand Est.

7.2.1.1. Multi-Camera Tracking in Partially Observable Environment

In collaboration with Thales ThereSIS - SE&SIM Team (Synthetic Environment & Simulation), we focus on the problem of following the trajectories of several persons with the help of several controllable cameras. This is a difficult problem since the set of cameras cannot simultaneously cover the whole environment, since some persons can be hidden by obstacles or by other persons, and since the behavior of each person is governed by internal variables which can only be inferred (such as his motivation or his hunger).

The approach we are working on is based on (1) the HMM (Hidden Markov Models) formalism to represent the state of the system (the persons and their internal states), (2) a simulator provided and developed by Thales ThereSIS, and (3) particle filtering approaches based on this simulator. Since activity and location depend on each other, we adopt a Simultaneous Tracking and Activity Recognition approach.

After having shown that it was possible to use a complex behavioral simulator to infer the behavior of complex individuals (motivation, possession, ...) even in case of long periods of occlusions [40], we investigated how to propose a factored particle filter (with one distribution per target) for efficiently tracking multiple targets simultaneously. To that end, we use a Joint Probabilistic Data Association Filter with a particular model of dynamics that largely decouples the evolution of several targets, and turns out to be very natural to apply. We proposed to use a small number of “representatives” of each target to determine and consider only effective interactions among targets.

This work has been published in Arsène Fansi Tchango’s PhD thesis which has been defended in December [7].

7.2.1.2. Human Posture Recognition

Human pose estimation in realistic world conditions raises multiple challenges such as foreground extraction, background update and occlusion by scene objects. Most of existing approaches were demonstrated in controlled environments. In this work, we propose a framework to improve the performance of existing tracking methods to cope with these problems. To this end, a robust and scalable framework is provided composed of three main stages. In the first one, a probabilistic occupancy grid updated with a Hidden Markov Model used to maintain an up-to-date background and to extract moving persons. The second stage uses component labelling to identify and track persons in the scene. The last stage uses a hierarchical particle filter to estimate the body pose for each moving person. Occlusions are handled by querying the occupancy grid to identify hidden body parts so that they can be discarded from the pose estimation process. We provide a parallel implementation that runs on CPU and GPU at 4 frames per second. We also validate the approach on our own dataset that consists of synchronized motion capture with a single RGB-D camera data of a person performing actions in challenging situations with severe occlusions generated by scene objects. We make this dataset available online (<http://www0.cs.ucl.ac.uk/staff/M.Firman/RGBDdatasets/>).

7.2.2. Social Robotics

Participants: Amine Boumaza, Serena Ivaldi.

We collaborate on this subject with Yann Boniface from Loria, Alain Dutech from Inria Nancy - Grand Est and Nicolas Rougier from the Mnemosyne team (Inria Bordeaux - Sud-Ouest).

7.2.2.1. PsyPhINE: Cogito Ergo Es

PsyPhINE is an interdisciplinary and exploratory project (see 9.1.2) between philosophers, psychologists and computer scientists. The goal of the project is related to cognition and behavior. Cognition is a set of processes that are difficult to unite in a general definition. The project aims to explore the idea of assignments of intelligence or intentionality, assuming that our intersubjectivity and our natural tendency to anthropomorphize play a central role: we project onto others parts of our own cognition. To test these hypotheses, our aim is to design a “non-verbal” Turing Test, which satisfies the definitions of our various fields (psychology, philosophy, neuroscience and computer science), using a robotic prototype. Some of the questions that we aim to answer are: is it possible to give the illusion of cognition and/or intelligence through such a technical device? How elaborate must be the control algorithms or “behaviors” of such a device so as to fool test subjects? How many degrees of freedom must it have?

Preliminary experiments with human subjects conducted this past year on a simple device helped to design an experimental protocol and test simple hypotheses which set the ground for the full fledged non verbal Turing Test. This project was funded under a PEPS Mirabelle grant (see 9.1.2) which helped build a robotic device with many degrees of freedom to perform further experiments. We also organized an inter-disciplinary workshop gathering top researchers from philosophy, anthropology, psychology and computer science to discuss and exchange on our methodology (see 10.1.1.1).

7.2.2.2. *Multimodal Object Learning During Human-Robot Interaction*

Robots working in evolving human environments need the ability to continuously learn to recognize new objects. Ideally, they should act as humans do, by observing their environment and interacting with objects, without specific supervision. However, if object recognition simply relies on visual input, then it may fail during human-robot interaction, because of the superposition of human and body parts. A multimodal approach was then proposed in [15], where visual input from cameras was combined with the robot proprioceptive information, in order to classify objects, robot, and human body parts. We proposed a developmental learning approach that enables a robot to progressively learn appearances of objects in a social environment: first only through observation, then through active object manipulation. We focused on incremental, continuous, and unsupervised learning that does not require prior knowledge about the environment or the robot. In the first phase of the proposed method, we analyse the visual space and detect proto-objects as units of attention that are learned and recognized as possible physical entities. The appearance of each entity is represented as a multi-view model based on complementary visual features. In the second phase, entities are classified into three categories: parts of the body of the robot, parts of a human partner, and manipulable objects. The categorization approach is based on mutual information between the visual and proprioceptive data, and on motion behaviour of entities. The ability to categorize entities is then used during interactive object exploration to improve the previously acquired objects models. The proposed system was implemented and evaluated with an iCub and a Meka robot learning 20 objects. The system was able to recognize objects with 88.5% success rate and create coherent representation models that are further improved by learning during human-robot interaction.

7.2.2.3. *Robot Functional and Social Acceptance*

To investigate the functional and social acceptance of a humanoid robot, we carried out an experimental study with 56 adult participants and the iCub robot. Trust in the robot has been considered as a main indicator of acceptance in decision-making tasks characterized by perceptual uncertainty (e.g., evaluating the weight of two objects) and socio-cognitive uncertainty (e.g., evaluating which is the most suitable item in a specific context), and measured by the participants' conformation to the iCub's answers to specific questions. In particular, we were interested in understanding whether specific (i) user-related features (i.e., desire for control), (ii) robot-related features (i.e., attitude towards social influence of robots), and (iii) context-related features (i.e., collaborative vs. competitive scenario), may influence their trust towards the iCub robot. We found that participants conformed more to the iCub's answers when their decisions were about functional issues than when they were about social issues. Moreover, the few participants conforming to the iCub's answers for social issues also conformed less for functional issues. Trust in the robot's functional savvy does not thus seem to be a pre-requisite for trust in its social savvy. Finally, desire for control, attitude towards social influence of robots and type of interaction scenario did not influence the trust in iCub. Results are also discussed with relation to methodology of HRI research in a currently submitted paper (<http://arxiv.org/abs/1510.03678> [cs.RO]). This work follows the research on engagement with social robots that was previously published [10].

7.2.2.4. *Relation Between Extroversion and Negative Attitude Towards Robot*

Estimating the engagement is critical for human - robot interaction. Engagement measures typically rely on the dynamics of the social signals exchanged by the partners, especially speech and gaze. However, the dynamics of these signals is likely to be influenced by individual and social factors, such as personality traits, as it is well documented that they critically influence how two humans interact with each other. We assess the influence of two factors, namely extroversion and negative attitude toward robots, on speech and gaze during a cooperative task, where a human must physically manipulate a robot to assemble an object [23]. We evaluate if the score of extroversion and negative attitude towards robots co-variate with the duration and frequency of gaze and

speech cues. The experiments were carried out with the humanoid robot iCub and 56 adult participants. We found that the more people are extrovert, the more and longer they tend to talk with the robot; and the more people have a negative attitude towards robots, the less they will look at the robot face and the more they will look at the robot hands where the assembly and the contacts occur. Our results confirm and provide evidence that the engagement models classically used in human-robot interaction should take into account attitudes and personality traits.

LEAR Project-Team

7. New Results

7.1. Visual recognition in images

7.1.1. Weakly Supervised Object Localization with Multi-fold Multiple Instance Learning

Participants: Ramazan Cinbis, Cordelia Schmid, Jakob Verbeek.

Object category localization is a challenging problem in computer vision. Standard supervised training requires bounding box annotations of object instances. This time-consuming annotation process is sidestepped in weakly supervised learning. In this case, the supervised information is restricted to binary labels that indicate the absence/presence of object instances in the image, without their locations. In [26], we propose to follow a multiple-instance learning approach that iteratively trains the detector and infers the object locations in the positive training images. Our main contribution is a multi-fold multiple instance learning procedure, which prevents training from prematurely locking onto erroneous object locations. Compared to state-of-the-art weakly supervised detectors, our approach better localizes objects in the training images, which translates into improved detection performance. Figure 1 illustrates the iterative object localization process on several example images. The technical report [26] is a journal paper under review after minor revision which extends a previous conference publication by adding experiments with CNN features, and a refinement procedure for the object location inference. These additions improve over related work that has appeared since the publication of the original paper.

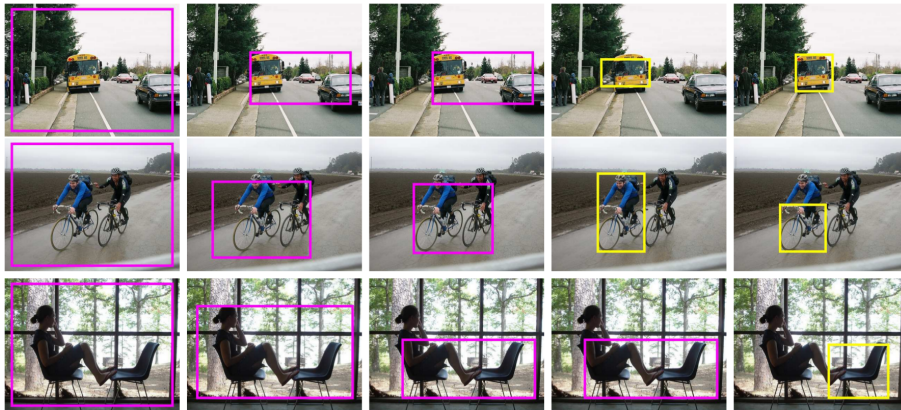


Figure 1. Illustration of our iterative object localization process on several example images, from initialization (left) to final localization (right). Yellow bounding boxes indicate that the object location hypothesis is in agreement with the ground-truth, for pink boxes the hypothesis is incorrect.

7.1.2. Patch-level spatial layout for classification and weakly supervised localization

Participants: Valentina Zadrija [University of Zagreb], Josip Krapac [University of Zagreb], Jakob Verbeek, Sinisa Segvic [University of Zagreb].

In [24] we propose a discriminative patch-level spatial layout model suitable for learning object localization models with weak supervision. We start from a block-sparse model of patch appearance based on the normalized Fisher vector representation. The appearance model is responsible for i) selecting a discriminative subset of visual words, and ii) identifying distinctive patches assigned to the selected subset. These patches are further filtered by a sparse spatial model operating on a novel representation of pairwise patch layout. We have evaluated the proposed pipeline in image classification and weakly supervised localization experiments on a public traffic sign dataset. The results show significant advantage of the proposed spatial model over state of the art appearance models.

7.1.3. *Approximate Fisher Kernels of non-iid Image Models for Image Categorization*

Participants: Ramazan Cinbis, Cordelia Schmid, Jakob Verbeek.

The bag-of-words (BoW) model treats images as sets of local descriptors and represents them by visual word histograms. The Fisher vector (FV) representation extends BoW, by considering the first and second order statistics of local descriptors. In both representations local descriptors are assumed to be identically and independently distributed (iid), which is a poor assumption from a modeling perspective. It has been experimentally observed that the performance of BoW and FV representations can be improved by employing discounting transformations such as power normalization. In [5], an expanded version of a previous conference publication, we introduce non-iid models by treating the model parameters as latent variables which are integrated out, rendering all local regions dependent. Using the Fisher kernel principle we encode an image by the gradient of the data log-likelihood w.r.t. the model hyper-parameters. Our models naturally generate discounting effects in the representations; suggesting that such transformations have proven successful because they closely correspond to the representations obtained for non-iid models. To enable tractable computation, we rely on variational free-energy bounds to learn the hyper-parameters and to compute approximate Fisher kernels. Our experimental evaluation results validate that our models lead to performance improvements comparable to using power normalization, as employed in state-of-the-art feature aggregation methods.

7.1.4. *Local Convolutional Features with Unsupervised Training for Image Retrieval*

Participants: Mattis Paulin, Matthijs Douze, Zaid Harchaoui, Julien Mairal, Florent Perronnin [Facebook], Cordelia Schmid.

Patch-level descriptors underlie several important computer vision tasks, such as stereo-matching or content-based image retrieval. We introduce a deep convolutional architecture that yields patch-level descriptors, as an alternative to the popular SIFT descriptor for image retrieval. The proposed family of descriptors, called Patch-CKN[17], adapt the recently introduced Convolutional Kernel Network (CKN), an unsupervised framework to learn convolutional architectures. We present a comparison framework to benchmark current deep convolutional approaches along with Patch-CKN for both patch and image retrieval (see Fig. 3 for our pipeline), including our novel “RomePatches” dataset. Patch-CKN descriptors yield competitive results compared to supervised CNNs alternatives on patch and image retrieval.

7.2. Learning and statistical models

7.2.1. *A Universal Catalyst for First-order Optimization*

Participants: Hongzhou Lin, Julien Mairal, Zaid Harchaoui.

In this paper [16], we introduce a generic scheme for accelerating first-order optimization methods in the sense of Nesterov, which builds upon a new analysis of the accelerated proximal point algorithm. Our approach consists of minimizing a convex objective by approximately solving a sequence of well-chosen auxiliary problems, leading to faster convergence. This strategy applies to a large class of algorithms, including gradient descent, block coordinate descent, SAG, SAGA, SDCA, SVRG, Finito/MISO, and their proximal variants. For all of these methods, we provide acceleration and explicit support for non-strongly convex objectives. In addition to theoretical speed-up, we also show that acceleration is useful in practice, as illustrated in Figure 4, especially for ill-conditioned problems where we measure significant improvements.

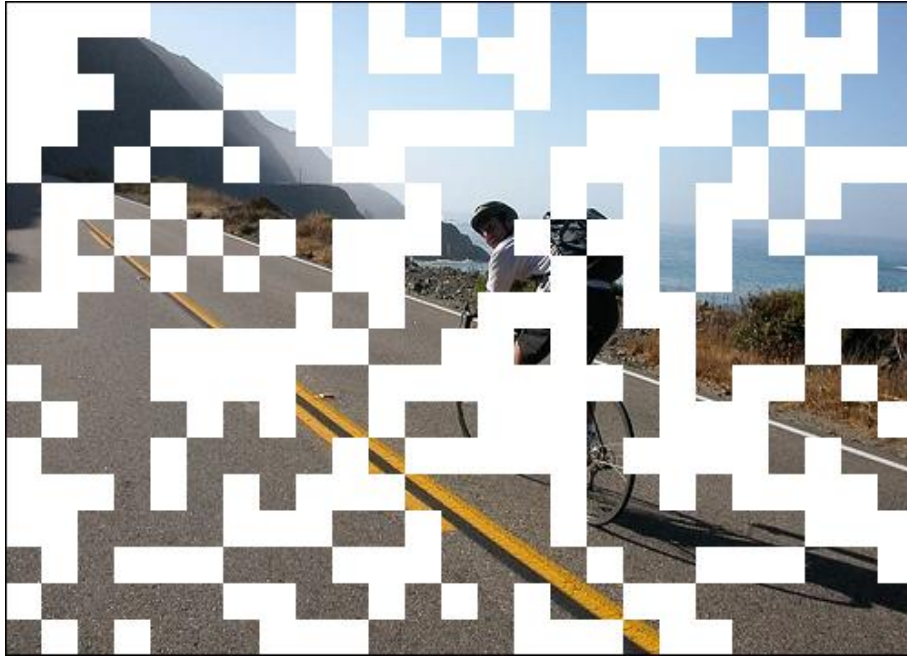


Figure 2. Illustration of why local image patches are not independent: we can easily guess the image content in the masked areas.

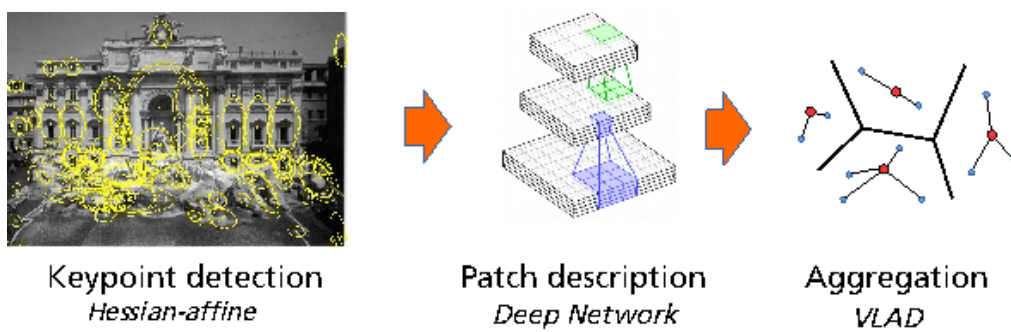


Figure 3. Image retrieval pipeline. Interest points are extracted with the Hessian-affine detector (left), encoded in descriptor space using convolutional features (middle), and aggregated into a compact representation using VLAD-pooling (right).

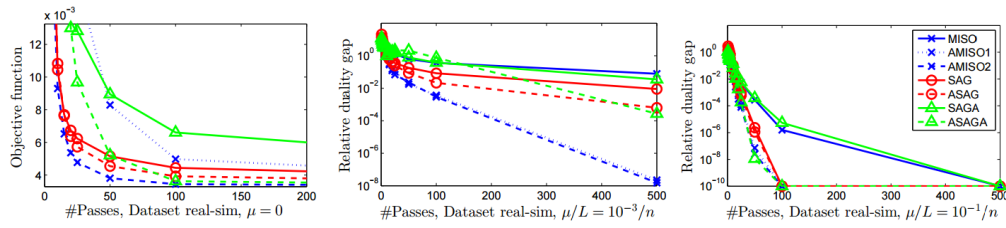


Figure 4. Objective function value (or duality gap) for different number of passes performed over each dataset. The legend for all curves is on the top right. AMISO, ASAGA, ASAG refer to the accelerated variants of MISO, SAGA, and SAG, respectively.

7.2.2. Incremental Majorization-Minimization Optimization with Application to Large-Scale Machine Learning

Participant: Julien Mairal.

In this paper [7], we study optimization methods consisting of iteratively minimizing surrogates of an objective function, as illustrated in Figure 5. We introduce a new incremental scheme that experimentally matches or outperforms state-of-the-art solvers for large-scale optimization problems typically arising in machine learning.

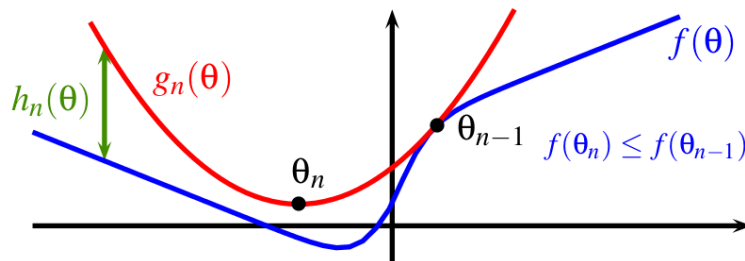


Figure 5. Illustration of the basic majorization-minimization principle. We compute a surrogate g_n of the objective function f around a current estimate θ_{n-1} . The new estimate θ_n is a minimizer of g_n . The approximation error h_n is smooth.

7.2.3. Coordinated Local Metric Learning

Participants: Shreyas Saxena, Jakob Verbeek.

Mahalanobis metric learning amounts to learning a linear data projection, after which the ℓ_2 metric is used to compute distances. In [20], we develop local metric learning techniques which allow more flexible metrics, not restricted to linear projections, see 6. Most of these methods partition the data space using clustering, and for each cluster a separate metric is learned. Using local metrics, however, it is not clear how to measure distances between data points assigned to different clusters. In this paper we propose to embed the local metrics in a global low-dimensional representation, in which the ℓ_2 metric can be used. With each cluster we associate a

linear mapping that projects the data to the global representation. This global representation directly allows computing distances between points regardless to which local cluster they belong. Moreover, it also enables data visualization in a single view, and the use of ℓ_2 -based efficient retrieval methods. Experiments on the Labeled Faces in the Wild dataset show that our approach improves over previous global and local metric learning approaches.

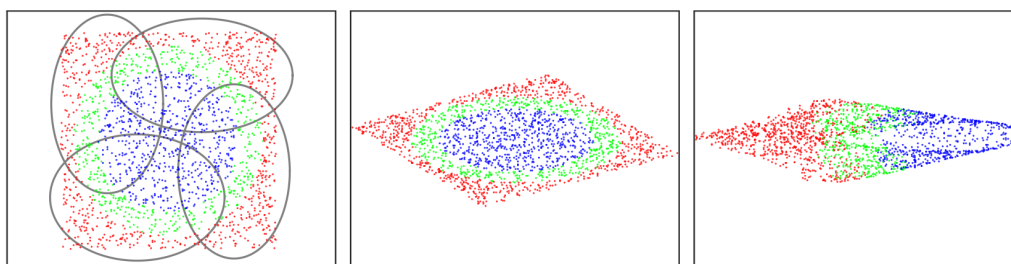


Figure 6. Synthetic dataset with color coded class labels, and the GMM used by our CLML local metric (left). Data projection given by a global Mahalanobis metric (middle) and our local CLML metric (right). The pairwise training constraints are better respected by CLML.

7.2.4. A convex formulation for joint RNA isoform detection and quantification from multiple RNA-seq samples

Participants: Elsa Bernard [Institut Curie, Ecoles des Mines-ParisTech], Laurent Jacob [CNRS, LBBE Laboratory], Julien Mairal, Jean-Philippe Vert [Institut Curie, Ecoles des Mines-ParisTech].

Detecting and quantifying isoforms from RNA-seq data is an important but challenging task. The problem is often ill-posed, particularly at low coverage. One promising direction is to exploit several samples simultaneously. In this paper [4], we propose a new method for solving the isoform deconvolution problem jointly across several samples. We formulate a convex optimization problem that allows to share information between samples and that we solve efficiently, as illustrated in Figure 7. We demonstrate the benefits of combining several samples on simulated and real data, and show that our approach outperforms pooling strategies and methods based on integer programming. Our convex formulation to jointly detect and quantify isoforms from RNA-seq data of multiple related samples is a computationally efficient approach to leverage the hypotheses that some isoforms are likely to be present in several samples. The software and source code are available at <http://cbio.ensmp.fr/flipflop>.

7.2.5. Adaptive Recovery of Signals by Convex Optimization

Participants: Zaid Harchaoui, Anatoli Juditsky [Univ. Grenoble], Arkadi Nemirovski [Georgia Tech], Dmitry Ostrovsky [Univ. Grenoble].

In [13], we present a theoretical framework for adaptive estimation and prediction of signals of unknown structure in the presence of noise. The framework allows to address two intertwined challenges: (i) designing optimal statistical estimators; (ii) designing efficient numerical algorithms. In particular, we establish oracle inequalities for the performance of adaptive procedures, which rely upon convex optimization and thus can be efficiently implemented. As an application of the proposed approach, we consider denoising of harmonic oscillations

7.2.6. Semi-proximal Mirror-Prox for Nonsmooth Composite Minimization

Participants: Niao He [Georgia Tech], Zaid Harchaoui.

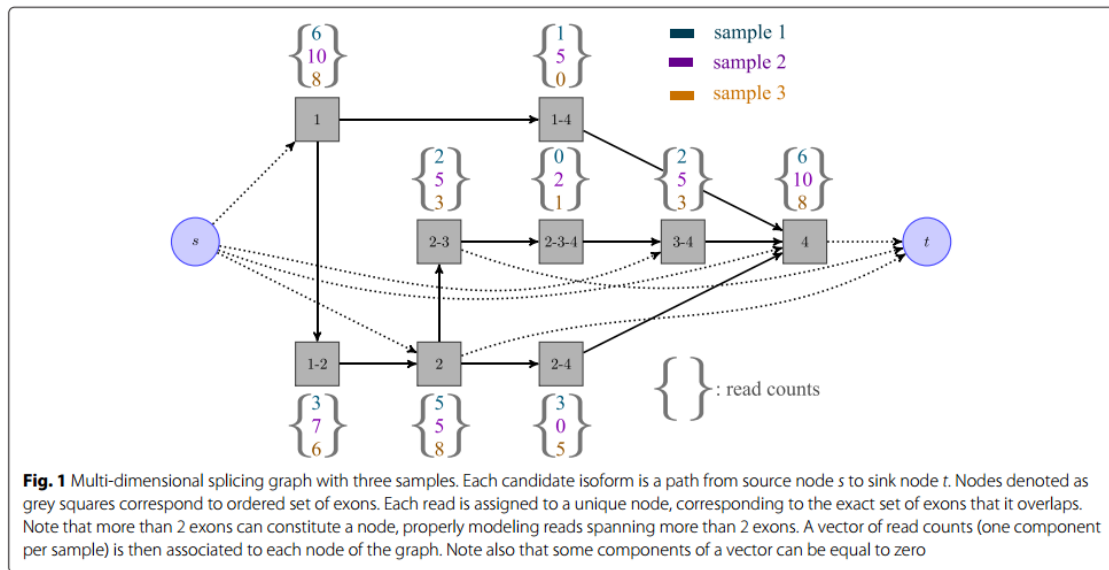


Figure 7. Graph on which we perform network flow optimization. Nodes represent observed reads, and paths on the graph correspond to isoforms.

In [28], we propose a new first-order optimisation algorithm to solve high-dimensional non-smooth composite minimisation problems. Typical examples of such problems have an objective that decomposes into a non-smooth empirical risk part and a non-smooth regularisation penalty. The proposed algorithm, called Semi-Proximal Mirror-Prox, leverages the Fenchel-type representation of one part of the objective while handling the other part of the objective via linear minimization over the domain. The algorithm stands in contrast with more classical proximal gradient algorithms with smoothing, which require the computation of proximal operators at each iteration and can therefore be impractical for high-dimensional problems. We establish the theoretical convergence rate of Semi-Proximal Mirror-Prox, which exhibits the optimal complexity bounds, for the number of calls to linear minimization oracle. We present promising experimental results showing the interest of the approach in comparison to competing methods.

7.3. Recognition in video

7.3.1. Beat-Event Detection in Action Movie Franchises

Participants: Danila Potapov, Matthijs Douze, Jerome Revaud, Zaid Harchaoui, Cordelia Schmid.

While important advances were recently made towards temporally localizing and recognizing specific human actions or activities in videos, efficient detection and classification of long video chunks belonging to semantically-defined categories such as “pursuit” or “romance” remains challenging.

In our work [30], we introduce a new dataset, Action Movie Franchises, consisting of a collection of Hollywood action movie franchises. We define 11 non-exclusive semantic categories — called beat-categories — that are broad enough to cover most of the movie footage. The corresponding beat-events are annotated as groups of video shots, possibly overlapping. We propose an approach for localizing beat-events based on classifying shots into beat-categories and learning the temporal constraints between shots, as shown in Figure 8. We show that temporal constraints significantly improve the classification performance. We set up an evaluation protocol for beat-event localization as well as for shot classification, depending on whether movies from the same franchise are present or not in the training data.



Figure 8. A 5-minute extract from the proposed Action Movie Franchises dataset, ground truth annotation and output of different methods. Each color stands for a different event category: green —pursuit, blue —battle, yellow —victory-good, green —despair-good, pink —romance, gray —victory-bad, cadet blue —good-argue-good. Hashes mark difficult examples. The color code for the classifier evaluation is: white = true positive, gray = ignored, black = false positive.

7.3.2. EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow

Participants: Jerome Revaud, Philippe Weinzaepfel, Zaid Harchaoui, Cordelia Schmid.

In this paper [18], we propose a novel approach for optical flow estimation, targeted at large displacements with significant occlusions. It consists of two steps: i) dense matching by edge-preserving interpolation from a sparse set of matches; ii) variational energy minimization initialized with the dense matches. The sparse-to-dense interpolation relies on an appropriate choice of the distance, namely an edge-aware geodesic distance. This distance is tailored to handle occlusions and motion boundaries – two common and difficult issues for optical flow computation. We also propose an approximation scheme for the geodesic distance to allow fast computation without loss of performance. Subsequent to the dense interpolation step, standard one-level variational energy minimization is carried out on the dense matches to obtain the final flow estimation. The proposed approach, called Edge-Preserving Interpolation of Correspondences (*EpicFlow*) is fast and robust to large displacements. An overview is given in Figure 9. EpicFlow significantly outperforms the state of the art on MPI-Sintel and performs on par on Kitti and Middlebury.

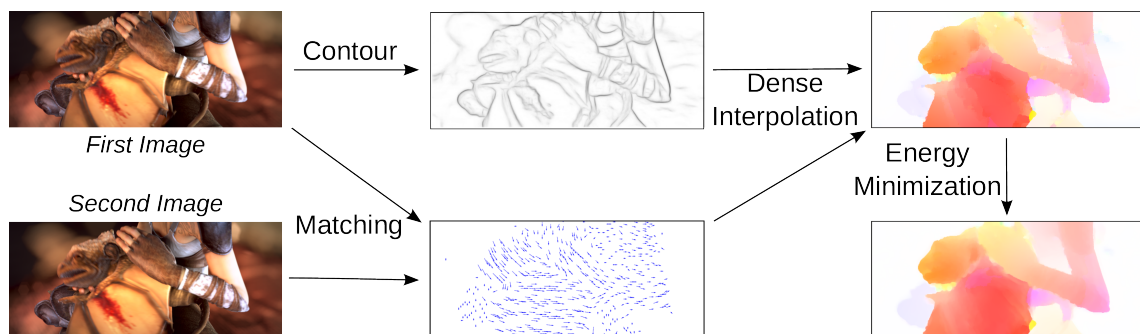


Figure 9. Overview of EpicFlow. Given two images, we compute matches using DeepMatching and the edges of the first image using SED. We combine these two cues to densely interpolate matches and obtain a dense correspondence field. This is used as initialization of a one-level energy minimization framework.

7.3.3. DeepMatching: Hierarchical Deformable Dense Matching

Participants: Jerome Revaud, Philippe Weinzaepfel, Zaid Harchaoui, Cordelia Schmid.

In this paper [31], we introduce a novel matching algorithm, called DeepMatching, to compute dense correspondences between images. DeepMatching relies on a hierarchical, multi-layer, correlational architecture designed for matching images and was inspired by deep convolutional approaches, see Figure 10. The proposed matching algorithm can handle non-rigid deformations and repetitive textures and efficiently determines dense correspondences in the presence of significant changes between images. We evaluate the performance of DeepMatching, in comparison with state-of-the-art matching algorithms, on the Mikolajczyk, the MPI-Sintel and the Kitti datasets. DeepMatching outperforms the state-of-the-art algorithms and shows excellent results in particular for repetitive textures. We also propose a method for estimating optical flow, called DeepFlow, by integrating DeepMatching in the large displacement optical flow (LDOF) approach of Brox et al. Compared to existing matching algorithms, additional robustness to large displacements and complex motion is obtained thanks to our matching approach. DeepFlow obtains competitive performance on public benchmarks for optical flow estimation.

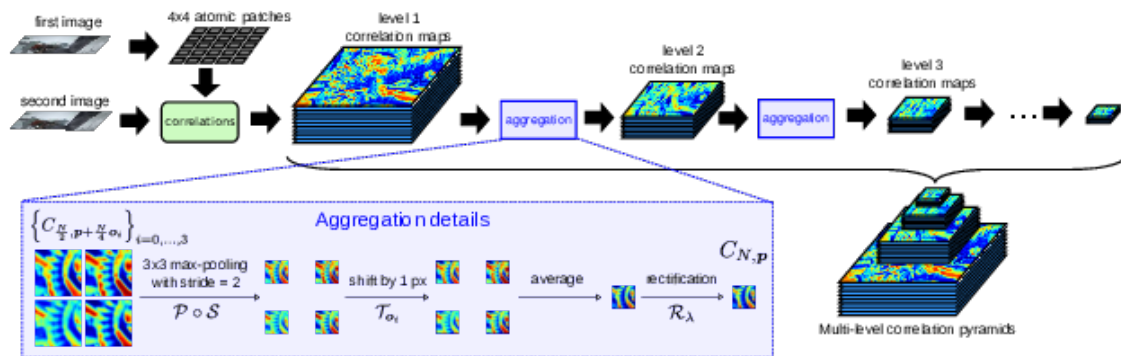


Figure 10. Overview of the bottom-up part of DeepMatching, which builds the multi-level correlation pyramid, from which matches are then extracted.

7.3.4. Learning to Detect Motion Boundaries

Participants: Philippe Weinzaepfel, Jerome Revaud, Zaid Harchaoui, Cordelia Schmid.

In this paper [23], we propose a learning-based approach for motion boundary detection. Precise localization of motion boundaries is essential for the success of optical flow estimation, as motion boundaries correspond to discontinuities of the optical flow field. The proposed approach allows to predict motion boundaries, using a structured random forest trained on the ground-truth of the MPI-Sintel dataset, see Figure 11. The random forest leverages several cues at the patch level, namely appearance (RGB color) and motion cues (optical flow estimated by state-of-the-art algorithms). Experimental results show that the proposed approach is both robust and computationally efficient. It significantly outperforms state-of-the-art motion-difference approaches on the MPI-Sintel and Middlebury datasets. We compare the results obtained with several state-of-the-art optical flow approaches and study the impact of the different cues used in the random forest. Furthermore, we introduce a new dataset, the YouTube Motion Boundaries dataset (YMB), that comprises 60 sequences taken from real-world videos with manually annotated motion boundaries. On this dataset, our approach, although trained on MPI-Sintel, also outperforms by a large margin state-of-the-art optical flow algorithms.

7.3.5. Learning to track for spatio-temporal action localization

Participants: Philippe Weinzaepfel, Zaid Harchaoui, Cordelia Schmid.

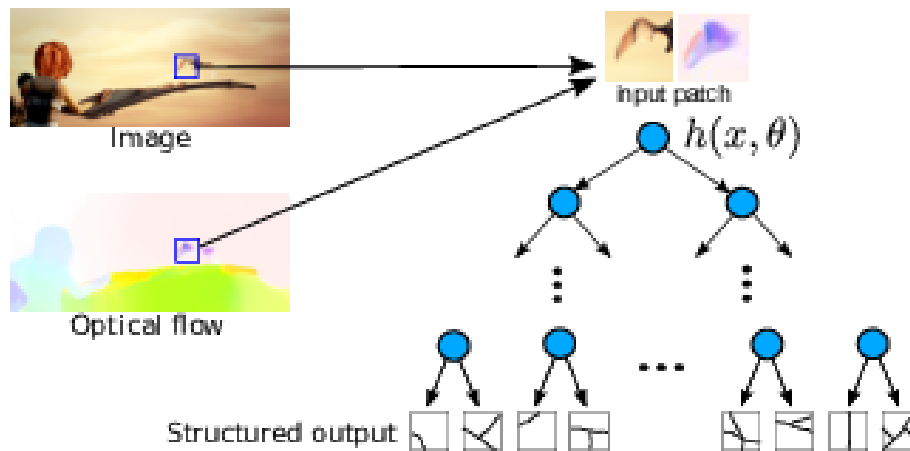


Figure 11. A structured random forest, taking as input a patch of an image and corresponding optical flow, outputs a motion boundaries patch. These motion boundaries patches are then aggregated to build the motion boundaries map for the whole image.

In this paper [22], we propose an effective approach for spatio-temporal action localization in realistic videos. The approach first detects proposals at the frame-level and scores them with a combination of static and motion CNN features. It then tracks high-scoring proposals throughout the video using a tracking-by-detection approach. Our tracker relies simultaneously on instance-level and class-level detectors. The tracks are scored using a spatio-temporal motion histogram, a descriptor at the track level, in combination with the CNN features. Finally, we perform temporal localization of the action using a sliding-window approach at the track level. An overview of our approach is given in Figure 12. We present experimental results for spatio-temporal localization on the UCF-Sports, J-HMDB and UCF-101 action localization datasets, where our approach outperforms the state of the art with a margin of 15%, 7% and 12% respectively in mAP.

7.3.6. A robust and efficient video representation for action recognition

Participants: Heng Wang, Dan Oneata, Cordelia Schmid, Jakob Verbeek.

In [9] we present a state-of-the-art video representation and apply it to efficient action recognition and detection. We first propose to improve the popular dense trajectory features by explicit camera motion estimation. Local feature trajectories consistent with the homography are considered as due to camera motion, and thus removed. This results in significant improvement on motion-based HOF and MBH descriptors. We further explore the recent Fisher vector as an alternative feature encoding approach to the standard bag-of-words histogram, and consider different ways to include spatial layout information in these encodings. We present a large and varied set of evaluations, considering (i) classification of short basic actions on six datasets, (ii) localization of such actions in featurelength movies, and (iii) large-scale recognition of complex events. We find that our improved trajectory features significantly outperform previous dense trajectories, and that Fisher vectors are superior to bag-of-words encodings for video recognition tasks. In all three tasks, we show substantial improvements over the state-of-the-art results. This journal paper combines and extends earlier conference papers.

7.3.7. Circulant temporal encoding for video retrieval and temporal alignment

Participants: Jerome Revaud, Matthijs Douze, Hervé Jégou [Inria Rennes, Facebook AI Research], Cordelia Schmid, Jakob Verbeek.

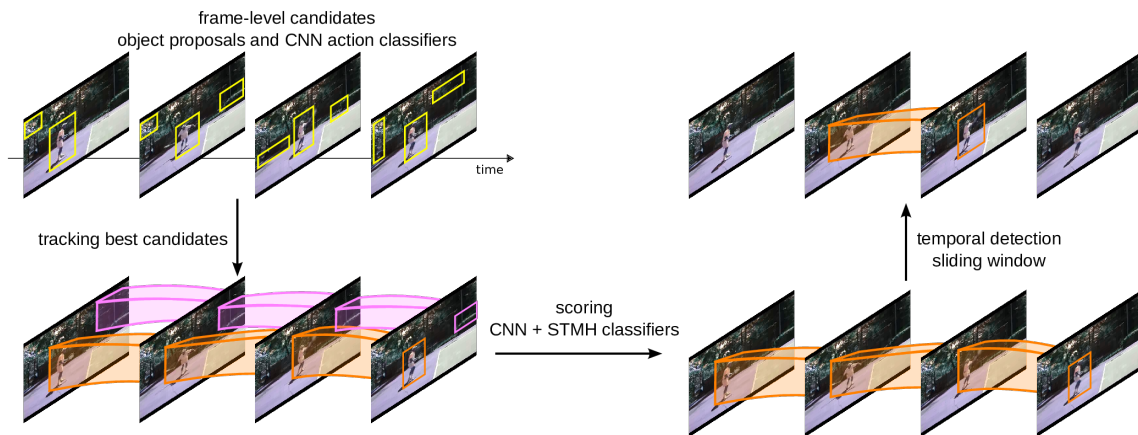


Figure 12. Overview of our spatio-temporal action localization approach. We detect frame-level object proposals and score them with CNN action classifiers. The best candidates, in term of scores, are tracked throughout the video. We then score the tracks with CNN and spatio-temporal motion histogram (STMH) classifiers. Finally, we perform a temporal sliding window for detecting the temporal extent of the action.

In [6] we address the problem of specific video event retrieval. Given a query video of a specific event, e.g., a concert of Madonna, the goal is to retrieve other videos of the same event that temporally overlap with the query. Our approach encodes the frame descriptors of a video to jointly represent their appearance and temporal order. It exploits the properties of circulant matrices to efficiently compare the videos in the frequency domain. This offers a significant gain in complexity and accurately localizes the matching parts of videos. The descriptors can be compressed in the frequency domain with a product quantizer adapted to complex numbers. In this case, video retrieval is performed without decompressing the descriptors. The second problem we consider is the temporal alignment of a set of videos. We exploit the matching confidence and an estimate of the temporal offset computed for all pairs of videos by our retrieval approach. Our robust algorithm aligns the videos on a global timeline by maximizing the set of temporally consistent matches. The global temporal alignment enables synchronous playback of the videos of a given scene. This journal paper extends an earlier conference paper.

7.3.8. Pose Estimation and Segmentation of Multiple People in Stereoscopic Movies

Participants: Guillaume Seguin [Willow], Karteek Alahari, Josef Sivic [Willow], Ivan Laptev [Willow].

The work in [8] presents a method to obtain a pixel-wise segmentation and pose estimation of multiple people in stereoscopic videos, as shown in Figure 13. This task involves challenges such as dealing with unconstrained stereoscopic video, non-stationary cameras, and complex indoor and outdoor dynamic scenes with multiple people. We cast the problem as a discrete labelling task involving multiple person labels, devise a suitable cost function, and optimize it efficiently. The contributions of our work are two-fold: First, we develop a segmentation model incorporating person detections and learnt articulated pose segmentation masks, as well as colour, motion, and stereo disparity cues. The model also explicitly represents depth ordering and occlusion. Second, we introduce a stereoscopic dataset with frames extracted from feature-length movies “StreetDance 3D” and “Pina”. The dataset contains 587 annotated human poses, 1158 bounding box annotations and 686 pixel-wise segmentations of people. The dataset is composed of indoor and outdoor scenes depicting multiple people with frequent occlusions. We demonstrate results on our new challenging dataset, as well as on the H2view dataset from (Sheasby et al. ACCV 2012).



Figure 13. We segment multiple people in the scene, estimate their poses and relative front-to-back order, denoted by the numbers in the image below, in every frame of a video sequence.

7.3.9. Encoding Feature Maps of CNNs for Action Recognition

Participants: Xiaojiang Peng, Cordelia Schmid.

In [29] We describe our approach for action classification in the THUMOS Challenge 2015. Our approach is based on two types of features, improved dense trajectories and CNN features, as illustrated in Figure 14. For trajectory features, we extract HOG, HOF, MBHx, and MBHy descriptors and apply Fisher vector encoding. For CNN features, we utilize a recent deep CNN model, VGG19, to capture appearance features and use VLAD encoding to encode/pool convolutional feature maps which shows better performance than average pooling of feature maps and full-connected activation features.

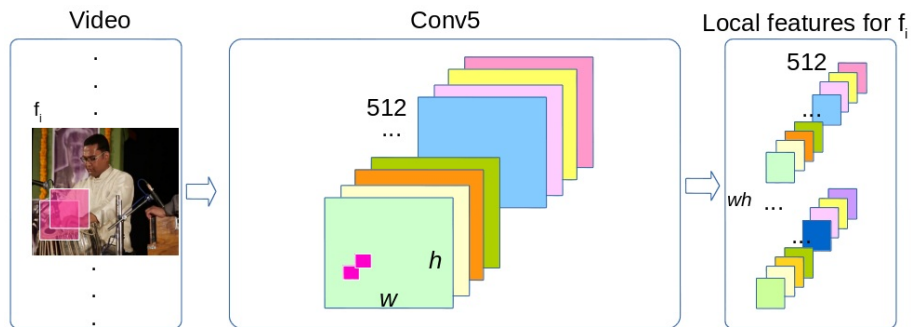


Figure 14. Local features from convolutional feature maps. Each pixel (pink square in the middle image) in the Conv5 feature map is actually a feature for the corresponding patch in original frame. We obtain $w \cdot h$ 512-D features for frame f_i .

7.3.10. Online Object Tracking with Proposal Selection

Participants: Yang Hua, Karteek Alahari, Cordelia Schmid.

Tracking-by-detection approaches are some of the most successful object trackers in recent years. Their success is largely determined by the detector model they learn initially and then update over time. However, under challenging conditions where an object can undergo transformations, e.g., severe rotation, these methods are found to be lacking. In [14], we address this problem by formulating it as a proposal selection task and making two contributions. The first one is introducing novel proposals estimated from the geometric transformations undergone by the object, and building a rich candidate set for predicting the object location. The second one is devising a novel selection strategy using multiple cues, i.e., detection score and edginess score computed from state-of-the-art object edges and motion boundaries. We extensively evaluate our approach on the visual object tracking 2014 challenge and online tracking benchmark datasets, and show the best performance. Sample results are shown in Figure 15. Our tracker based on this method has recently won the visual object tracking challenge (VOT-TIR) organized as part of ICCV 2015 in Santiago, Chile.

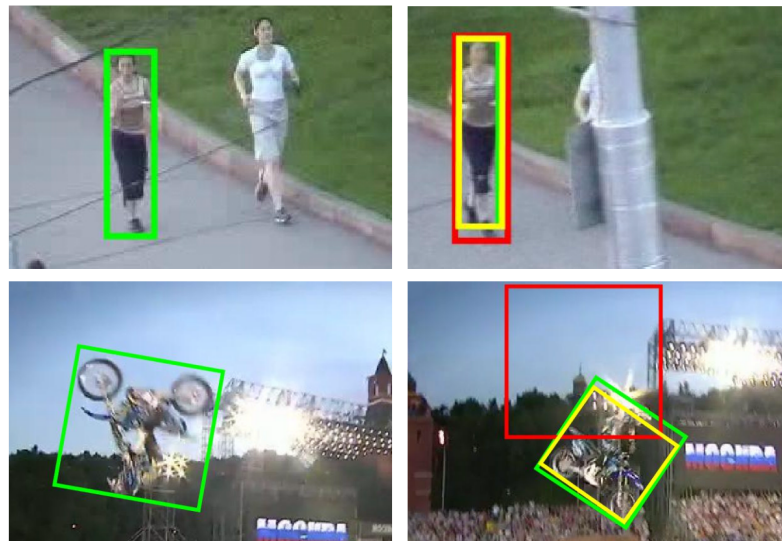


Figure 15. Sample frames (cropped) from the jogging (top row) and motocross (bottom row) sequences. The ground truth annotation (green) in the first frame (left) is used to train our tracker and the winner of VOT2014 challenge. We show these two tracking results (right) on another frame in the sequence. Our method (yellow) successfully tracks objects undergoing deformations unlike winner of VOT2014 challenge (red).

LINKMEDIA Project-Team

6. New Results

6.1. Unsupervised motif and knowledge discovery

6.1.1. Estimation of continuous intrinsic dimension

Participants: Laurent Amsaleg, Teddy Furon.

In collaboration with Michael Houle, National Institute for Informatics (Japan).

Some of our research work was concerned with the estimation of continuous intrinsic dimension (ID), a measure of intrinsic dimensionality recently proposed by Houle. Continuous ID can be regarded as an extension of Karger and Ruhl's expansion dimension to a statistical setting in which the distribution of distances to a query point is modeled in terms of a continuous random variable. This form of intrinsic dimensionality can be particularly useful in search, classification, outlier detection, and other contexts in machine learning, databases, and data mining, as it has been shown to be equivalent to a measure of the discriminative power of similarity functions. In [11], we proposed several estimators of continuous ID that we analyzed based on extreme value theory, using maximum likelihood estimation, the method of moments, probability weighted moments, and regularly varying functions. Experimental evaluation was performed using both real and artificial data.

6.1.2. Supervised multi-scale locality sensitive hashing

Participants: Laurent Amsaleg, Li Weng.

LSH is a popular framework to generate compact representations of multimedia data, which can be used for content based search. However, the performance of LSH is limited by its unsupervised nature and the underlying feature scale. In [42], we proposed to improve LSH by incorporating two elements: supervised hash bit selection and multi-scale feature representation. First, a feature vector is represented by multiple scales. At each scale, the feature vector is divided into segments. The size of a segment is decreased gradually to make the representation correspond to a coarse-to-fine view of the feature. Then each segment is hashed to generate more bits than the target hash length. Finally the best ones are selected from the hash bit pool according to the notion of bit reliability, which is estimated by bit-level hypothesis testing. Extensive experiments have been performed to validate the proposal in two applications: near-duplicate image detection and approximate feature distance estimation. We first demonstrate that the feature scale can influence performance, which is often a neglected factor. Then we show that the proposed supervision method is effective. In particular, the performance increases with the size of the hash bit pool. Finally, the two elements are put together. The integrated scheme exhibits further improved performance.

6.1.3. Rotation and translation covariant match kernels for image retrieval

Participants: Andrei Bursuc, Teddy Furon, Hervé Jégou, Giorgos Tolias.

Most image encodings achieve orientation invariance by aligning the patches to their dominant orientations and translation invariance by completely ignoring patch position or by max-pooling. Albeit successful, such choices introduce too much invariance because they do not guarantee that the patches are rotated or translated consistently. In this work, we propose a geometric-aware aggregation strategy, which jointly encodes the local descriptors together with their patch dominant angle [38] and/or location [10]. The geometric attributes are encoded in a continuous manner by leveraging explicit feature maps. Our technique is compatible with generic match kernel formulation and can be employed along with several popular encoding methods, in particular bag of words, VLAD and the Fisher vector. The method is further combined with an efficient monomial embedding to provide a codebook-free method aggregating local descriptors into a single vector representation. Invariance is achieved by efficient similarity estimation of multiple rotations or translations, offered by a simple trigonometric polynomial. This strategy is effective for image search, as shown by experiments performed on standard benchmarks for image and particular object retrieval, namely Holidays and Oxford buildings.

6.1.4. Sequential pattern mining on audio data

Participants: Laurent Amsaleg, Guillaume Gravier, Simon Malinowski.

M. Sc. Internship of Corentin Hardy, in collaboration with René Quiniou, Inria Rennes, DREAM research team, within the framework of the STIC AmSud Maximum project and of the MOTIF Inria Associate Team.

Analyzing multimedia data is a challenging problem due to the quantity and complexity of such data. Mining for frequently recurring patterns is a task often ran to help discovering the underlying structure hidden in the data. This year, we have explored how data symbolization and sequential pattern mining techniques could help for mining recurring patterns in multimedia data. In [20], we have shown that even if sequential pattern mining techniques are very helpful in terms of computational efficiency, the data symbolization step is a crucial step to find for extracting relevant audio patterns.

6.1.5. Clustering by diverting supervised machine learning

Participants: Vincent Claveau, Teddy Furon, Guillaume Gravier.

M. Sc. Internship of Amélie Royer, ENS Rennes.

Clustering algorithms exploit an input similarity measure on the samples, which should be fine-tuned with the data format and the application at hand. However, manually defining a suitable similarity measure is a difficult task in case of limited prior knowledge or complex data structures for example. While supervised classification systems require a set of samples annotated with their ground-truth classes, recent studies have shown it is possible to exploit classifiers trained on an artificial annotation of the data in order to induce a similarity measure. In this work, we have proposed a unified framework, named similarity by iterative classifications (SIC), which explores the idea of diverting supervised learning for automatic similarity inference. We studied several of its theoretical and practical aspects. We also have implemented and evaluate SIC on three tasks of knowledge discovery on multimedia content. Results show that in most situations the proposed approach indeed benefits from the underlying classifier's properties and outperforms usual similarity measures for clustering applications.

6.1.6. Multimodal person discovery in TV broadcasts

Participant: Guillaume Gravier.

Work in collaboration with Cassio Elias dos Santos Jr. and William Robson Schwartz, in the framework of the Inria Associate Team MOTIF and of the STIC AmSud project Maximum.

Taking advantage of recent results on large-scale face comparison with partial least square, we developed various approaches for multimodal person discovery in TV broadcasts in the framework of the MediaEval 2015 international benchmark [30]. The task consists in naming the persons on screen that are speaking with no prior information, leveraging text overlays, speech transcripts as well as face and voice comparison. We investigated two distinct aspects of multimodal person discovery. One refers to face clusters, which are considered to propagate names associated with faces in one shot to other faces that probably belong to the same person. The face clustering approach consists in calculating face similarities using partial least squares and a simple hierarchical approach. The other aspect refers to tag propagation in a graph-based approach where nodes are speaking faces and edges link similar faces/speakers. The advantage of the graph-based tag propagation is to not rely on face/speaker clustering, which we believe can be errorprone. The face clustering approach ranked among the top results in the international benchmark.

6.1.7. Unsupervised video structure mining with grammatical inference

Participants: Guillaume Gravier, Bingqing Qu.

In collaboration with Jean Carrive and Félicien Vallet, Institut National de l'Audiovisuel.

In [25], we addressed the problem of unsupervised program structuring with minimal prior knowledge about the program. We extended previous work to propose an approach able to identify multiple structures and infer structural grammars for recurrent TV programs of different types. The approach taken involves three sub-problems: i) we determine the structural elements contained in programs with minimal knowledge about which type of elements may be present; ii) we identify multiple structure for the programs if any and model the structures of programs; iii) we generate the structural grammar for each corresponding structure. Finally, we conducted use-case based evaluations on real recurrent programs of three different types to demonstrate the effectiveness of the proposed approach.

6.1.8. Information retrieval for distributional semantics, and vice-versa

Participants: Vincent Claveau, Ewa Kijak.

Distributional thesauri are useful in many tasks of natural language processing. In [33], [3], we address the problem of building and evaluating such thesauri with the help of information retrieval (IR) concepts. Two main contributions are proposed. First, in the continuation of previous work, we have shown how IR tools and concepts can be used with success to build thesauri. Through several experiments and by evaluating directly the results with reference lexicons, we show that some IR models outperform state-of-the-art systems. Secondly, we use IR as an application framework to indirectly evaluate the generated thesaurus. Here again, this task-based evaluation validate the IR approach used to build the thesaurus. Moreover, it allows us to compare these results with those from the direct evaluation framework used in the literature. The observed differences question these evaluation habits.

6.2. Multimedia content description and structuring

6.2.1. Image description using component trees

Participants: Petra Bosilj, Ewa Kijak.

In collaboration with Sébastien Lefèvre from Obelix Team (IRISA).

In this work, we explored the application of a tree-based feature extraction algorithm for the widely-used MSER features, and proposed a tree-of-shapes based detector of maximally stable regions. Changing an underlying component tree in the algorithm allows considering alternative properties and pixel orderings for extracting maximally stable regions. Performance evaluation was carried out on a standard benchmark in terms of repeatability and matching score under different image transformations, as well as in a large scale image retrieval setup, measuring mean average precision. The detector outperformed the baseline MSER in the retrieval experiments [37].

We also proposed a local region descriptor based on 2D shape-size pattern spectra, calculated on arbitrary connected regions, and combined with normalized central moments. The challenges when transitioning from global pattern spectra to the local ones were faced, and an exhaustive study on the parameters and the properties of the newly constructed descriptor was conducted. The descriptors were calculated on MSER regions, and evaluated in a simple retrieval system. Competitive performance with SIFT descriptors was achieved. An additional advantage of the proposed descriptors is their size which is less than half the size of SIFT [14], [15].

6.2.2. Improved motion description for action classification

Participant: Hervé Jégou.

In collaboration with Mihir Jain (University of Amsterdam, The Netherlands) and Patrick Bouthemy (Team-project SERPICO, Inria Rennes, France)

Even though the importance of explicitly integrating motion characteristics in video descriptions has been demonstrated by several recent papers on action classification, our current work concludes that adequately decomposing visual motion into dominant and residual motions, i.e., camera and scene motion, significantly improves action recognition algorithms. This holds true both for the extraction of the space-time trajectories and for computation of descriptors. We designed in [7] a new motion descriptor—the DCS descriptor—that captures additional information on local motion patterns enhancing results based on differential motion scalar quantities, divergence, curl and shear features. Finally, applying the recent VLAD coding technique proposed in image retrieval provides a substantial improvement for action recognition. These findings are complementary to each other and they outperformed all previously reported results by a significant margin on three challenging datasets: Hollywood 2, HMDB51 and Olympic Sports as reported in (Jain et al. (2013)).

6.2.3. Word embeddings and recurrent neural networks for spoken language understanding

Participants: Guillaume Gravier, Christian Raymond, Vedran Vukotić.

Recently, word embedding representations have been investigated for slot filling in spoken language understanding (SLU), along with the use of neural networks as classifiers. Neural networks, especially recurrent neural networks, which are adapted to sequence labeling problems, have been applied successfully on the popular ATIS database. In [29], we make a comparison of this kind of models with the previously state-of-the-art conditional random fields (CRF) classifier on a more challenging SLU database. We show that, despite efficient word representations used within these neural networks, their ability to process sequences is still significantly lower than for CRF, while also having a drawback of higher computational costs, and that the ability of CRF to model output label dependencies is crucial for SLU.

6.2.4. Hierarchical topic structuring

Participants: Guillaume Gravier, Pascale Sébillot, Anca-Roxana Şimon.

Topic segmentation traditionally relies on lexical cohesion measured through word re-occurrences to output a dense segmentation, either linear or hierarchical. We have proposed a novel organization of the topical structure of textual content [28]. Rather than searching for topic shifts to yield dense segmentation, our algorithm extracts topically focused fragments organized in a hierarchical manner. This is achieved by leveraging the temporal distribution of word re-occurrences, searching for bursts, to skirt the limits imposed by a global counting of lexical re-occurrences within segments. Comparison to a reference dense segmentation on varied datasets indicates that we can achieve a better topic focus while retrieving all of the important aspects of a text.

6.2.5. Partial least square hashing for large-scale face identification

Participants: Guillaume Gravier, Ewa Kijak.

Work performed with Cassio Elias dos Santos Jr. during his 3 months visit, in collaboration with William Robson Schwartz (UFMG, Brasil), in the framework of the Inria Associate Team MOTIF.

Face recognition has been largely studied in past years. However, most of the related work focus on increasing accuracy and/or speed to test a single pair probe-subject. In [31], we introduced a novel method inspired by the success of locality sensing hashing applied to large general purpose datasets and by the robustness provided by partial least squares analysis when applied to large sets of feature vectors for face recognition. The result is a robust hashing method compatible with feature combination for fast computation of a short list of candidates in a large gallery of subjects. We provided theoretical support and practical principles for the proposed hashing method that may be reused in further development of hash functions applied to face galleries. Comparative evaluations on the FERET and FRGCv1 datasets demonstrate a speedup of a factor 16 compared to scanning all subjects in the face gallery.

6.2.6. Selection strategies for active learning in NLP

Participants: Vincent Claveau, Ewa Kijak.

Nowadays, many NLP problems are modeled as supervised machine learning tasks, especially when it comes to information extraction. Consequently, the cost of the expertise needed to annotate the examples is a widespread issue. Active learning offers a framework to that issue, allowing to control the annotation cost while maximizing the classifier performance, but it relies on the key step of choosing which example will be proposed to the expert. In [3], we have examined and proposed such selection strategies in the specific case of conditional random fields which are largely used in NLP. On the one hand, we have proposed a simple method to correct a bias of certain state-of-the-art selection techniques. On the other hand, we have detailed an original approach to select the examples, based on the respect of proportions in the datasets. These contributions are validated over a large range of experiments implying several tasks and datasets, including named entity recognition, chunking, phonetization, word sense disambiguation.

6.2.7. *Tree-structured named entities extraction from competing speech transcripts*

Participant: Christian Raymond.

When real applications are working with automatic speech transcription, the first source of error does not originate from the incoherence in the analysis of the application but from the noise in the automatic transcriptions. In [41], we present a simple but effective method to generate a new transcription of better quality by combining utterances from competing transcriptions. We have extended a structured named entity (NE) recognizer submitted during the ETAPE challenge. Working on French TV and radio programs, our system revises the transcriptions provided by making use of the NEs it has detected. Our results suggest that combining the transcribed utterances which optimize the F-measure, rather than minimizing the WER scores, allows the generation of a better transcription for NE extraction. The results show a small but significant improvement of 0.9 % SER against the baseline system on the ROVER transcription. These are the best performances reported to date on this corpus.

6.3. Content-based information retrieval

6.3.1. *A comparison of dense region detectors for image search and fine-grained classification*

Participants: Hervé Jégou, Ahmet Iscen, Giorgos Tolias.

In collaboration with Philippe-Henri Gosselin (ETIS team, ENSEA, Cergy, France)

We consider a pipeline for image classification or search based on coding approaches like bag of words or Fisher vectors. In this context, the most common approach is to extract the image patches regularly in a dense manner on several scales. In [6], we propose and evaluate alternative choices to extract patches densely. Beyond simple strategies derived from regular interest region detectors, we propose approaches based on super-pixels, edges, and a bank of Zernike filters used as detectors. The different approaches are evaluated on recent image retrieval and fine-grain classification benchmarks. Our results show that the regular dense detector is outperformed by other methods in most situations, leading us to improve the state of the art in comparable setups on standard retrieval and fine-grain benchmarks. As a byproduct of our study, we show that existing methods for blob and super-pixel extraction achieve high accuracy if the patches are extracted along the edges and not around the detected regions.

6.3.2. *Efficient large-scale similarity search using matrix factorization*

Participants: Teddy Furon, Ahmet Iscen.

In collaboration with Michael Rabbat (McGill University, Montréal, Canada)

We considered the image retrieval problem of finding the images in a dataset that are most similar to a query image. Our goal is to reduce the number of vector operations and memory for performing a search without sacrificing accuracy of the returned images. We adopt a group testing formulation and design the decoding architecture using either dictionary learning or eigendecomposition. The latter is a plausible option for small-to-medium sized problems with high-dimensional global image descriptors, whereas dictionary learning is applicable in large-scale scenario. We evaluate our approach both for global descriptors obtained from SIFT and CNN features. Experiments with standard image search benchmarks, including the Yahoo100M dataset

comprising 100 million images, show that our method gives comparable (and sometimes superior) accuracy compared to exhaustive search while requiring only 10 % of the vector operations and memory. Moreover, for the same search complexity, our method gives significantly better accuracy compared to approaches based on dimensionality reduction or locality sensitive hashing [43].

6.3.3. *Explicit embeddings for nearest neighbor search with Mercer kernels*

Participant: Hervé Jégou.

In collaboration with Anthony Bourrier and Patrick Pérez (Technicolor, Rennes, France), Florent Perronnin (Xerox, Grenoble, France) Rémi Gribonval (Team-project PANAMA, Inria Rennes, France).

Many approximate nearest neighbor search algorithms operate under memory constraints, by computing short signatures for database vectors while roughly keeping the neighborhoods for the distance of interest. Encoding procedures designed for the Euclidean distance have attracted much attention in the last decade. In the case where the distance of interest is based on a Mercer kernel, we propose a simple, yet effective two-step encoding scheme: first, compute an explicit embedding to map the initial space into a Euclidean space; second, apply an encoding step designed to work with the Euclidean distance. Comparing this simple baseline with existing methods relying on implicit encoding, we demonstrate better search recall for similar code sizes with the chi-square kernel in databases comprised of visual descriptors, outperforming concurrent state-of-the-art techniques by a large margin [2].

6.3.4. *Image search with selective match kernels: aggregation across single and multiple images*

Participants: Hervé Jégou, Giorgos Tolias.

In collaboration with Yannis Avrithis (National Technical University of Athens, Greece)

Our work [9] considers a family of metrics to compare images based on their local descriptors. It encompasses the VLAD descriptor and matching techniques such as Hamming Embedding. Making the bridge between these approaches leads us to propose a match kernel that takes the best of existing techniques by combining an aggregation procedure with a selective match kernel. The representation underpinning this kernel is approximated, providing a large scale image search both precise and scalable, as shown by our experiments on several benchmarks. We show that the same aggregation procedure, originally applied per image, can effectively operate on groups of similar features found across multiple images. This method implicitly performs feature set augmentation, while enjoying savings in memory requirements at the same time. Finally, the proposed method is shown effective for place recognition, outperforming state of the art methods on a large scale landmark recognition benchmark.

6.3.5. *Early burst detection for memory-efficient image retrieval*

Participant: Hervé Jégou.

In collaboration with Miaojing Shi, visiting Ph. D. student from Peking University, and Yannis Avrithis (National Technical University of Athens, Greece)

Recent works show that image comparison based on local descriptors is corrupted by visual bursts, which tend to dominate the image similarity. The existing strategies, like power-law normalization, improve the results by discounting the contribution of visual bursts to the image similarity. We proposed to explicitly detect the visual bursts in an image at an early stage. We compare several detection strategies jointly taking into account feature similarity and geometrical quantities. The bursty groups are merged into meta-features, which are used as input to state-of-the-art image search systems such as VLAD or the selective match kernel. Then, we show the interest of using this strategy in an asymmetrical manner, with only the database features being aggregated but not those of the query. Extensive experiments performed on public benchmarks for visual retrieval show the benefits of our method, which achieves performance on par with the state of the art but with a significantly reduced complexity, thanks to the lower number of features fed to the indexing system [40], [44].

6.3.6. Biomedical information retrieval

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with N. Grabar (STL), T. Hamon (LIMSI), and S. Le Maguer (Univ. Saarland).

The right of patients to access their clinical health record is granted by the code of Santé Publique. Yet, this piece of content remains difficult to understand. We propose different IR experiments in which we use queries defined by patients in order to find relevant documents [3], [16]. We use the Indri search engine, based on statistical language modeling, as well as semantic resources. More precisely, our approaches are chiefly based on the terminological variation (e.g., synonyms, abbreviations) to link between expert and patient languages. Various combinations of resources and Indri settings are explored, mostly based on query expansion.

6.4. Linking, navigation and analytics

6.4.1. Sentiment analysis on social networks

Participants: Vincent Claveau, Christian Raymond, Vedran Vukotić.

In the framework of our participation to the DeFT 2015 text-mining challenge, we have developed sentiment-analysis methods for tweets [34]. Several sub-tasks have been considered: i) valence classification of tweets and ii) fine-grained classification of tweets (which includes two sub-tasks: detection of the generic class of the information expressed in a tweet and detection of the specific class of the opinion/sentiment/emotion. For all three problems, we adopt a standard machine learning framework. More precisely, three main methods are proposed and their feasibility for the tasks is analyzed: i) decision trees with boosting (bonzaiboost), ii) naive Bayes with Okapi and iii) convolutional neural networks (CNNs). Our approaches are voluntarily knowledge free and text-based only, we do not exploit external resources (lexicons, corpora) or tweet metadata. It allows us to evaluate the interest of each method and of traditional bag-of-words representations vs. word embeddings. Methods using simple ML frameworks and IR-based similarity metrics have been demonstrated to yield the best results.

6.4.2. A multi-dimensional data model for personal photo browsing

Participant: Laurent Amsaleg.

Work performed in the framework of the CNRS PICS MMAAnalytics, and in collaboration with Marcel Worring, Univeristy of Amsterdam (The Netherlands)

Digital photo collections—personal, professional, or social—have been growing ever larger, leaving users overwhelmed. It is therefore increasingly important to provide effective browsing tools for photo collections. Learning from the resounding success of multi-dimensional analysis (MDA) in the business intelligence community for on-line analytical processing (OLAP) applications, we proposed a multi-dimensional model for media browsing, called M3, that combines MDA concepts with concepts from faceted browsing [21]. We present the data model and describe preliminary evaluations, made using server and client prototypes, which indicate that users find the model useful and easy to use.

6.4.3. NLP-driven hyperlink construction in broadcast videos

Participants: Rémi Bois, Guillaume Gravier, Pascale Sébillot, Anca-Roxana Şimon.

In collaboration with Sien Moens (Katholieke Universiteit Leuven, Belgium), Éric Jamet and Martin Ragot (Univ. Rennes 2, France).

In the context of the the CominLabs project "Linking media in acceptable hypergraphs" dedicated to the creation of explicit and meaningful links between multimedia documents or fragments of documents, we have introduced a typology of possible links between contents of a multimedia news corpus [32]. While several typologies have been proposed and used by the community, we argue that they are not adapted to rich and large corpora which can contain texts, videos, or radio stations recordings. We have defined a new typology, as a first step towards automatically creating and categorizing links between documents' fragments in order to create new ways to navigate, explore, and extract knowledge from large collections.

We also investigated video hyperlinking based on speech transcripts, leveraging a hierarchical topical structure to address two essential aspects of hyperlinking, namely, serendipity control and link justification [26]. We proposed and compared different approaches exploiting a hierarchy of topic models as an intermediate representation to compare the transcripts of video segments. These hierarchical representations offer a basis to characterize the hyperlinks, thanks to the knowledge of the topics which contributed to the creation of the links, and to control serendipity by choosing to give more weights to either general or specific topics. Experiments have been performed on BBC videos from the Search and Hyperlinking task at MediaEval. Link precisions similar to those of direct text comparison have been achieved however exhibiting different targets along with a potential control of serendipity.

The Search and Anchoring in Video Archives task at MediaEval addressed two issues: The Search part aims at returning a ranked list of video segments that are relevant to a textual user query; The Anchoring part focuses on identifying video segments that would encourage further exploration within the archive. Capitalizing on the experience acquired in previous participations, we implemented a two step approach for both sub-tasks [27]. The first step, common to both, consists in generating a list of potential anchor segments and response-query segments relying on a hierarchical topical structuring technique. In the second step, for each query, the best 20 segments are selected according to content-based comparisons, while for the anchor detection sub-task, the segments are ranked based on a cohesion measure. The use of a hierarchical topical structure helps to propose segments of variable length at different levels of details with precise jump-in points for them. More, the algorithm deriving the structure relies on the burstiness phenomenon in word occurrences which gives an advantage over the classical bag-of-words model.

6.4.4. Information extraction

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with X. Tannier (LIMSI), A. Vilnat (LIMSI) and B. Arnulphy (ANR).

Identifying events from texts is an information extraction task necessary for many NLP applications. Through the TimeML specifications and TempEval challenges, it has received some attention in the last years; yet, no reference result is available for French. In [12], we try to fill this gap by proposing several event extraction systems, combining for instance Conditional Random Fields, language modeling and k-nearest-neighbors. These systems are evaluated on French corpora and compared with state-of-the-art methods on English. The very good results obtained on both languages validate our whole approach.

6.5. Participation in benchmarking initiatives

- Video hyperlinking, TRECVID
- Search and anchoring, Mediaeval Multimedia International Benchmark
- Multimodal person discovery in broadcast TV, Mediaeval Multimedia International Benchmark
- DeFT 2015 text-mining challenge

LINKS Team

7. New Results

7.1. Querying Heterogeneous Linked Data

7.1.1. Recursive queries

P. Bourhis published a paper at IJCAI [17] in cooperation with the University of Dresden in Germany. There he developed a highly expressive Web query language of the Datalog family, for which static analysis problems such as query containment remain decidable.

In cooperation with Links' associated team in Oxford, P. Bourhis obtained an article at ACM TODS [5], where he studies the access of hidden data by recursive queries.

V. Hugot, A. Boiret, and J. Niehren study monadic second-order logic for unordered trees with data constraints on siblings. This language can be used to define recursive queries and schemas on unordered data trees [13]. They study restrictions of the logics, for which the usual static analysis problems become decidable, and study the complexity of the decidable cases. This work was done in cooperation with Paris 7.

7.1.2. Schemas

I. Boneva and S. Staworko contribute at ICDT the RDF schema language SheX [22], which they developed in cooperation with members of the W3C. The usual open world approach of RDF is schemaless in that the alphabets of RDF data are left open, so that data from different sources and with different alphabets can be unified. This raises serious problems for query writing and thus for linked data integration, since a query may become invalid when the alphabet changes. A SheX schema allows to express constraints on the alphabets, node labels and edge labels of RDF graphs, so that database queries become safe with respect to future changes without closing the alphabet. In a previous work the studied the case of XML data trees instead of RDF graphs [6].

A. Lemay and J. Niehren propose sublinear algorithms in the style of probabilistic property testing for validating XML data trees with respect to DTD [20].

P. Bourhis studies streaming bounded repair with respect to schema violations [8]. This work is done in a cooperation with the University of Bordeaux and the University of Santiago in Chile.

7.1.3. Provenance

P. Bourhis obtained an ICALP paper [11] in cooperation with Télécom ParisTech. They show how to propagate provenance information for monadic second-order logics on trees or tree like structures with polynomial data complexity. In their provenance framework, they can show how to generalize various aggregation tasks for monadic second-order logics, that were known to be solvable with polynomial data complexity before.

In a cooperation with Tel Aviv, P. Bourhis obtained a ACM CIKM paper [18], where they show how to approximately summarize data provenance.

7.1.4. Data integration

In a cooperation with the University of Toronto, R. Ciucanu obtained a paper at PVLDB [4] on how to gain control over data integration evaluations. I. Boneva, A. Bonifati and R. Ciucaniu presented a paper on graph data exchange with target constraints [14] in the GraphQ workshop, and proved that query answering is intractable in this context.

7.2. Managing Dynamic Linked Data

7.2.1. *Complex event processing*

T. Sebastian, J. Niehren and D. Debarbieux propose early nested word automata for evaluating navigational XPath queries on XML streams [9]. They show how to approximate earliest query answering for such queries in a highly efficient manner and with very good precision in practice, while exact earliest query answering is known to untractable for XPath. This work was done in an industrial cooperation with Innovimax from Paris and in cooperation with the University of Bordeaux. In a follow-up work [21] they show that the XPath streaming algorithm for early nested word automata can be speed up considerably, when combining it with projection algorithms for nested word automata that they developed.

J. Niehren developed X-Fun [19] a uniform programming language for implementing XML standards, and showed how to implement XSLT, XProc, and XSLT in this manner. This work, that is fully implemented, was done in cooperation with the University of Bratislava.

7.2.2. *Data-centered workflows*

P. Bourhis presents highly expressive query languages as needed for data-centric workflows in the context of Active XML [3] in cooperation with the Dahu project from Inria Saclay.

J. Niehren presents a general framework for the reasoning with observational program semantics [10] in a cooperation with the Universities of Frankfurt and Saarbrücken in Germany.

7.3. Linking Data Graphs

S. Staworko obtained his HDR for his work on symbolic inference methods for databases [2]. R. Ciucanu obtained his PhD for his work on cross-model query inference [1] supervised by A. Bonifati.

7.3.1. *Learning path queries*

A. Lemay, R. Ciucanu, and A. Bonifati have a paper and a demo at EDBT showing how to learn simple path queries on graph databases based on automata techniques [16], [15], [25], [24]. This is a very interesting starting point for using automata inference techniques in the context of graph databases.

S. Staworko obtained a paper at ICDT where he shows how to infer XML Twig queries from examples [23]. This work is done in cooperation with the University of Wrazlaw.

7.3.2. *Learning join queries*

R. Ciucanu, A. Boneva, and S. Staworko published an ACM TODS article [7] showing how to learn join queries for relational databases from examples. This is the first query learning algorithm satisfying Gold's learning model, that relies on equalities of data values rather than on structural information.

MAGNET Team

7. New Results

7.1. Decentralized Estimation in Networks

In [3], we studied the problem of decentralized estimation in networks, where each node of the network holds a data point and the goal is to estimate some statistics on the entire data under communication constraints imposed by the graph topology of the network. This generic problem has many applications in Internet of Things as well as for extracting knowledge from massive information graphs such as interlinked Web documents and online social media. In this work, we focused on estimating pairwise mean statistics. Popular examples of such statistics include the sample variance, the average distance and the Area Under the ROC Curve, among others. We proposed new synchronous and asynchronous randomized gossip algorithms which simultaneously propagate data across the network and maintain local estimates of the quantity of interest. We establish convergence rate bounds of $O(1/t)$ and $O(\log t/t)$ for the synchronous and asynchronous cases respectively, where t is the number of iterations, with explicit data and network dependent terms. Beyond favorable comparisons in terms of rate analysis, numerical experiments provide empirical evidence the proposed algorithms surpasses the previously introduced approach.

7.2. Large-Scale Learning with Higher-Order Risk Functionals

In [6], we studied learning problems where the performance criterion consists of an average over tuples (e.g., pairs or triplets) of observations rather than over individual observations, as in many learning problems involving networked data (e.g., link prediction), but also in metric learning and ranking. In this setting, the empirical risk to be optimized takes the form of a U -statistic, and its terms are highly dependent and thus violate the classic i.i.d. assumption. In this work, we focused on how to best implement a stochastic approximation approach to solve such risk minimization problems in the large-scale setting. We argue that gradient estimates should be obtained by sampling tuples of data points with replacement (incomplete U -statistics) rather than sampling data points without replacement (complete U -statistics based on subsamples). We develop a theoretical framework accounting for the substantial impact of this strategy on the generalization ability of the prediction model returned by the Stochastic Gradient Descent (SGD) algorithm. It reveals that the method we promote achieves a much better trade-off between statistical accuracy and computational cost. Beyond the rate bound analysis, we provide strong empirical evidence of the superiority of the proposed approach on metric learning and ranking problems.

7.3. Natural Language Processing

In [4], we introduce a new structured model for learning anaphoricity detection and coreference resolution in a joint fashion. Specifically, we use a latent tree to represent the full coreference and anaphoric structure of a document at a global level, and we jointly learn the parameters of the two models using a version of the structured perceptron algorithm. Our joint structured model is further refined by the use of pairwise constraints which help the model to capture accurately certain patterns of coreference. Our experiments on the CoNLL-2012 English datasets show large improvements in both coreference resolution and anaphoricity detection, compared to various competing architectures. Our best coreference system obtains a CoNLL score of 81.97 on gold mentions, which is to date the best score reported on this setting.

In [2], we present a detailed comparative framework for assessing the usefulness of unsupervised word representations for identifying so-called implicit discourse relations. Specifically, we compare standard one-hot word pair representations against low-dimensional ones based on Brown clusters and word embeddings. We also consider various word vector combination schemes for deriving discourse segment representations from word vectors, and compare representations based either on all words or limited to head words. Our main finding is that denser representations systematically outperform sparser ones and give state-of-the-art performance or above without the need for additional hand-crafted features.

7.4. Some Ongoing Work

7.4.1. Metric Learning for Graph-based Label Propagation

The efficiency of graph-based semi-supervised algorithms depends on the graph of instances on which they are applied. The instances are often in a vectorial form before a graph linking them is built. The construction of the graph relies on a metric over the vectorial space that helps define the weight of the connection between entities. The typical choice for this metric is usually a distance or a similarity measure based on the Euclidean norm. We claim that in some cases the Euclidean norm on the initial vectorial space might not be the most appropriate to solve the task efficiently.

In a paper currently under review, we proposed an algorithm that aims at learning the most appropriate vectorial representation for building a graph on which label propagation is solved efficiently, with theoretical guarantees on the classification performance.

7.4.2. Link Classification in Signed Graphs

We worked on active link classification in signed graphs. Namely, the idea is to build a spanning tree of the graph and query all its edge signs. In the two clusters case, this allows to predict the sign of an edge between nodes u and v as the product of the signs of edge along the path in the spanning tree from u to v . It turns out that ensuring low error rate amounts to minimizing the stretch, a long open standing problem known as Low Stretch Spanning Tree [11]. While we are still working on the theoretical analysis, experimental results showed that our construction is generally competitive with a simple yet efficient baseline and outperforms it for specific graph geometry like grid graphs.

Moreover, based on experimental observations, we will also analyze a heuristic which exhibits good performance at a very low computational cost and is therefore well suited for large-scale graphs. In a nutshell, it predicts the sign of an edge from u to v based on the fraction of u negative outgoing edges and v negative incoming edges, exploiting a behavioral consistency bias from signed social network users.

Going further in link classification, we believe that the notion of sign can be extended, going from one binary label per edge to a more holistic approach where the similarity between two nodes is measured across different contexts. These contexts are represented by vectors whose dimension matches the dimension of unknown feature vectors associated with each node. The goal is to answer queries of the form: how similar are nodes u and v along a specific context? We first plan to validate the relevance of this modeling on real-world problems, then test baseline methods on synthetic and real data before looking for a more effective, online prediction method.

7.4.3. Graph-based Learning for Dependency Parsing

We are investigating the use of different graph-based learning techniques such as k -nearest neighbors classification and label propagation for the problem of dependency parsing. While most of current approaches rely on learning a single scoring model (through SVM, MIRA, neural networks) from a large set of hand annotated training data (usually thousands of sentences), we are interested in using the sentence space geometry (approximated via a similarity graph over some labeled and unlabeled sentences) to tune the model to better fit a given sentence. This amounts to learning a slightly different model for each unlabeled sentence.

In order to successfully parse sentences in this setting, we need to propagate parsing information from labeled sentences to unlabeled ones through the graph. In order to build a similarity graph well suited to dependency parsing, we worked on learning a similarity function between pairs of sentences, based on the idea that two sentences are similar if they have similar parse trees. We will then investigate how to propagate the trees (which may be of varying sizes) through the graph and consider several propagation schemes.

MAGRIT Project-Team

6. New Results

6.1. Matching and 3D tracking

Participants: Marie-Odile Berger, Jaime Garcia Guevara, Nazim Haouchine, Gilles Simon, Frédéric Sur.

Pose initialization Automating the camera pose initialization is still a problem in non instrumented environments. Difficulties originate in the possibly large viewpoint changes between the data stored in the model and the current view. In this context, Pierre Rolin's PhD work concerns viewpoint simulation techniques for localization. The idea is to generate keypoint descriptors from simulated views in order to enrich the model and to ease the matching of the current view to the model. We have demonstrated the effectiveness of this technique in several situations, either under an affine or a perspective camera model [17], [21]. The computed pose is more stable when it is difficult to obtain reliable correspondences between the model and the current view. In addition, several examples show that our method successfully computes the camera pose whereas the traditional methods fail. Our recent work concerns a progressive sampling strategy to speed up the search of correspondences when confronted to a large outlier rate, which is inherent to viewpoint simulation. We also currently investigate the localization of the virtual camera from which viewpoints should be simulated.

AR in urban environments

Pose initialization is especially difficult in urban scenes due to the presence of repeated patterns. Another difficulty originates in the fact that a pedestrian is free of his motion in the scene and can therefore adopt uncontrolled viewpoints (close or distant views) with respect to the model. As a result, the set of 2D/3D correspondence hypotheses may contain a high ratio of outliers which may lead to erroneous pose computation. In order to improve the matching / recognition stage, we investigated how facades in calibrated images can be orthorectified and delimited by considering prior information about the scene and the camera relevant to AR in urban context [20]. We provide a Bayesian framework to detect vanishing points in Manhattan worlds, which incorporate priors about the Manhattan frame by imposing a near-vertical direction as well as orthogonality constraints. Second, we propose to detect right-angle corners due to windows or doors using a SVM-based machine learning technique. Rectangular facade hypotheses are then generated through min-cuts techniques with the idea to identify rectangles with high density of right-angle corners. Our algorithm performs better or as well as state-of-the-art techniques and is much faster, mainly as a result of using a suitable prior.

Tracking 3D deformable objets

3D augmentation of deformable objects is a challenging problem with many potential applications in computer graphics, augmented reality and medical imaging. Most existing approaches are dedicated to surface augmentation and are based on the inextensibility constraint, for sheet-like materials, or on the use of a model built from representative samples. However, few of them consider in-depth augmentation which is of utmost importance for medical applications. Since the beginning of N. Haouchine's PhD thesis, we have addressed several important limitations that currently hinder the use of augmented reality in the clinical routine of minimally invasive procedures. In collaboration with the MIMESIS team, our main contribution is the design and the validation of an augmented reality framework based on a mechanical model of the organ and guided by features extracted and tracked on the video at the surface of the organ [12]. Specific models which best suit the considered organs, such as a vascularized model of the liver, have been introduced in this framework. Experiments conducted on ex-vivo data of a porcine liver show that the localization error of a virtual tumor were less than 6mm, and thus below the safety margin required by surgery. To our knowledge, we were the first to produce such evaluation for deformable objects.

This work has been extended to augment highly elastic objects in a monocular context. Shape recovery from a monocular video sequence is an underconstrained problem. State-of-the art solutions enforce smoothness or geometric constraints, consider specific deformation properties such as inextensibility or resort to shading constraints. However, few of them can handle properly large elastic deformations. We have proposed [13] a real-time method that uses a mechanical model and is able to handle highly elastic objects. The problem is formulated as an energy minimization problem accounting for a non-linear elastic model constrained by external image points acquired from a monocular camera. This method prevents us from formulating restrictive assumptions and specific constraint terms in the minimization. In addition, we propose to handle self-occluded regions thanks to the ability of mechanical models to provide appropriate predictions of the shape.

The work conducted during N. Haouchine's PhD thesis allowed us to build a complete framework for the use of AR in liver surgery. We now want to focus on specific points to improve the accuracy and the robustness of the augmented process and to facilitate the clinical use of such AR systems. The PhD thesis of Jaime Garcia Guevara started in October on this topic with the aim to build more realistic mechanical models of organs during the surgery (taking into account liver deformation due to insufflation of air during surgery) and to improve the robustness of visual tracking through the use of multiple visual cues and improved methods for outlier detection.

6.2. Image-based modeling

Participants: Marie-Odile Berger, Charlotte Delmas, Antoine Fond, Erwan Kerrien, Gilles Simon, Pierre-Frédéric Villard, Brigitte Wrobel-Dautcourt.

Finding Manhattan directions in uncalibrated images

Finding orthogonal vanishing points (VPs) in a photography has many potential applications in computer vision and computer graphics, including perspective correction, image-based reconstruction and texture extraction. Surprisingly, while this problem has been extensively studied in the literature, manual solutions are still used in most software. Existing algorithms generally follow two steps. First, lines are grouped into pencils, whose centers are considered as potential VPs. Then, an orthogonality measure is evaluated for every triplet of VPs and the most plausible triplet is used for camera calibration. A drawback of this approach is that complex and time-consuming techniques have to be used to solve the general problem of VP detection, while only three particular VPs are finally used. By contrast, we propose an effective and easy-to-implement algorithm that estimates the zenith and the horizon line before detecting the VPs, using simple properties of the central projection and exploiting accumulations of oriented segments around the horizon. Our method is fast and yields an accuracy comparable, and even better in some cases, to that of state-of-the-art algorithms. This work was submitted to Eurographics 2016.

Tools reconstruction for interventional neuro-radiology

Minimally invasive techniques impact surgery in such ways that, in particular, an imaging modality is required to maintain a visual feedback. Live X-ray imaging, called fluoroscopy, is used in interventional neuroradiology. Such images are very noisy, and cannot show but the vasculature and no other brain tissue. In particular, since at most only two projective fluoroscopic views are available, containing absolutely no depth hint, the 3D shape of the micro-tool (guidewire, micro-catheter or micro-coil) can be very difficult, if not impossible to infer, which may have an impact on the clinical outcome of the procedure.

In collaboration with GE Healthcare, we aim at devising ways to reconstruct the micro-tools in 3D from fluoroscopy images. Charlotte Delmas has been working as a PhD CIFRE student on this subject since April 2013. She first devised a solution in a two-view reconstruction context. A paper reporting on this work was published and an oral presentation was made at SPIE Medical Imaging 2015 [19]. The focus was made this year on reconstructing the first coil, a single wire that tangles into a complex pattern when deployed in an aneurysm. The challenge is to produce a 3D reconstruction with as little X-ray dose and acquisition time as possible. Two paths were simultaneously followed this year: 1) design and compare various configurations to rapidly shoot 6 X-ray views from different viewpoints with a biplane (stereo) system; 2) compensate the lack of information (small number of images) with a prior, incorporated in the tomographic reconstruction algorithm, to express the sparsity in space of the shape to be reconstructed. Preliminary work sets forward a

simultaneous fast rotation of both planes around the patient's head. A database is currently being acquired for a full validation in a view to submitting this work for publication early next year.

Individual-specific heart valve modeling

Mitral valve surgical repair is a complex procedure where the outcome largely depends on the surgeons' experience. Predicting a good coaptation of the two leaflets post-operatively is one of the most difficult sub-tasks in the procedure. We worked on a biomechanical simulation framework [25] that computes the leaflet deformation and coaptation based on individual-specific microtomography data as an initial step toward patient-based mitral valve surgical repair assistance through simulation. Results from FEM analysis on three explanted porcine hearts showed that it is possible to obtain the real shape of the leaflets during systolic peak. We also measured the influence of the positions of chordae tendineae on the simulation and showed that marginal chordae have a greater influence on the final shape than intermediary chordae.

Quasi-periodic noise removal

Images may be affected by quasi-periodic noise. This undesirable feature manifests itself by spurious repetitive patterns covering the whole image, well localized in the Fourier domain. While notch filtering permits to get rid of this phenomenon, this however requires to first detect the resulting Fourier spikes, and, in particular, to discriminate between noise spikes and spectrum patterns caused by spatially localized textures or repetitive structures. Several approaches have been investigated. First, we have reviewed the available methods, most of them requiring expert tuning, with applications to experimental mechanics in view [11]. We have also proposed two automated patch-based approaches. A parametric approach has been investigated in [14] (more information available in [26]), based on the detection of noise spikes as statistical outliers to the distribution expected from natural non-noisy patches, which is known to follow the inverse of a power of the frequency. A non-parametric approach, based on a-contrario detection, was also proposed in [22].

6.3. Parameter estimation

Participants: Frédéric Sur, Erwan Kerrien, Raffaella Trivisonne.

Metrological performance assessment in experimental mechanics

A problem of interest in experimental solid mechanics is to estimate displacement and strain maps on the surface of a specimen subjected to a load or a tensile test. Two contactless approaches are available in the literature, based on depositing on the surface of the specimen either a pseudo-periodic grid or a random speckle. Analyzing images taken before and after deformation permits to estimate strain maps. While periodicity permits to make use of Fourier analysis in the first case, digital image correlation (DIC) is used in the second case. Concerning pseudo-periodic grids, we have investigated noise reduction techniques. While averaging a series of images is certainly the most basic option to reduce the noise, it is impossible to get rid of residual vibrations carried for instance by concrete floor slabs. We have shown in [16] that, while these vibrations indeed blur grid images, they still permit to reduce the noise amplitude in the displacement and strain maps. Concerning DIC-based techniques, we have investigated the effect of sensor noise on the measurement resolution. Since displacement of interest are most of the time below one pixel, interpolation plays a major role. We have proposed a new resolution formula which takes interpolation into account. Besides, this formula has been the subject of an experimental assessment on real data, in the presence of signal-dependent noise [24], [18].

Sensor noise measurement

We have investigated in [15] (additional information available in [27]) the problem of sensor parameter estimation from a series of images, under illumination flickering and vibrations. Illumination flickering is indeed a natural assumption for indoor artificial lights. It is also involved by slight variations in the opening time of a mechanical shutter. We have proposed a model of the pixel intensity based on a Cox process, together with an algorithm which, taking benefit of flickering, gives an estimation of every sensor parameter, namely the gain, the readout noise, and the offset.

Image driven simulation

The IDeaS ANR project targets image-driven simulation, applied to interventional neuroradiology: a coupled system of interactive computer-based simulation (interventional devices in blood vessels) and on-line medical image acquisitions (X-ray fluoroscopy). The main idea is to use the live X-ray images as references to continuously refine the parameters used to simulate the blood vessels and the interventional devices (micro-guide, micro-catheter, coil). Our guideline is to follow a sequential statistical filtering approach to fuse such heterogeneous data.

Christo Gnonnou was hired as an engineer (located at Inria Lille in Defrost team (ex-Shacra), contract started on January 1st and ended on October 31st). He continued the work to specify which parameters the simulation is sensitive to, in a view to design a reduced parametric model of the device, and associate covariances to its parameters. He also worked on inverting the mechanical parameters of any device, using our experimental setup based on a high speed stereo rig.

Maxime Malgras worked on his Master's thesis in the team. His investigations aimed at designing a particle filter where the transition function is approximated by a polynomial chaos (PC) instead of particles. It appeared that PC is adapted to compute the posterior in a particle filter but it is not clear whether the number of samples required to estimate the PC coefficients is smaller than the number of particles required for the filter to be accurate, which questions the capacity of PC to reduce the computational burden of particle filters in high dimensions. Raffaella Trivisonne started her PhD thesis in November (co-supervised by Stéphane Cotin, from MIMESIS team in Strasbourg) to investigate deeper on this subject and apply data assimilation to image driven simulation of endovascular interventions.

MANAO Project-Team

7. New Results

7.1. Analysis and Simulation

7.1.1. Parametrization of BRDFs

Opaque materials are represented in computer graphics by Bi-directional Reflectance Distribution Functions (BRDF), which are 4D functions of light and view direction. Dealing with such a high dimensionality is problematic for the modeling and rendering of material appearance. The choice of a BRDF parametrization greatly simplifies this task by identifying the axis where most variations occur in common opaque materials. The 4D parametrization of Rusinkiewicz [86] is classically used in graphics, in particular because of its direct connection to micro-facet theory. Alternative parametrization by Neumann et al. [71] and Stark et al. [91] have been proposed, but are restricted to 2D parametrizations, and hence a restricted class of materials.

We have extended the work of Neumann et al. [71] and Stark et al. [91] to a pair of 4D BRDF parameterizations with explicit changes of variables. Revealing some of their mathematical properties and relationships to Rusinkiewicz' parametrization allows us to better understand their benefits and drawbacks for representing measured BRDFs. Our preliminary study suggests that the alternative parametrization inspired by Stark et al. [91] is superior, and should thus be considered in future work involving BRDFs.

7.1.2. New BRDF Model and Diffraction Effects identification

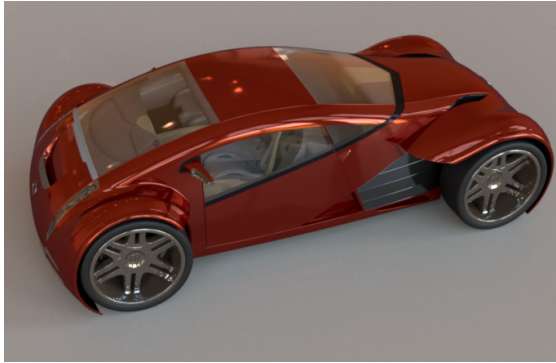
Finding the appropriate BRDF model, with meaningful physical parameters, that can represent accurately measured data remains a challenging task. In [20], we show that two different physical phenomena are present in measured reflectance: reflection and diffraction. Taking both into account, we present a reflectance model [24] that is compact and a very good approximation (cf. Figure 8) of measured reflectance. Designers can act on model parameters, related to surface properties, to create new materials.

7.1.3. Statistical analysis of BRDFs

On the one hand, a BRDF is a complex 4D function, which should ensure reciprocity and energy conservation laws. On the other hand, when computing radiance reaching the eye from a surface point, the view direction is held fixed. In this respect, we are only interested in a 2D BRDF slice that acts as a filter on the local environment lighting. In [21], our goal is to understand the statistical properties of such a filter as a function of viewing elevation. To this end, we have conducted a study of measured BRDFs where we have computed statistical moments for each viewing angle. We show that some moments are correlated together across dimensions and orders, while some others are close to zero and may safely be discarded. Our study opens the way to novel applications such as moment-based manipulation of measured BRDFs, material estimation and image-based material editing. It also puts empirical and physically-based material models in a new perspective, by revealing their effect as view-dependent filters.

7.1.4. Importance Sampling of Real Light Sources

Realistic images can be rendered by simulating light transport with Monte Carlo methods. The possibility to use realistic light sources for synthesizing images greatly contributes to their physical realism. Among existing models, the ones based on light fields are attractive due to their ability to capture faithfully the far-field and near-field effects as well as their possibility of being acquired directly. Since acquired light sources have arbitrary frequencies and possibly high dimensions (4D), using such light sources for realistic rendering leads to performance problems. We have investigated [12] how to balance the accuracy of the representation and the efficiency of the simulation (cf. Figure 9). The work relies on generating high quality samples from the input light sources for unbiased Monte Carlo estimation [67]. This is a foundation work that has led to new sampling techniques for physically-based rendering with light field light sources. The results show that physically accurate rendering with realistic light sources can be achieved in real time.



- (a) Car model (Path tracing with adaptive sampling, 128 to 16384 samples per pixel), with smoother red-metallic-paint body, chrome wheels, white-marble floor. Inside the car: pickled-oak-260, specular-white-phenolic
- (b) Kitchen model (Energy Redistribution Path Tracing, 5h), with colonial-maple-223 cupboards, chrome tap, sink and handles, nickel kitchen wall, alumina-oxide oven door, white-marble and black-obsidian tiles, brass bowls, aluminium glasses.

Figure 8. Example scenes using our new BRDF model [24] to represent different measured materials.

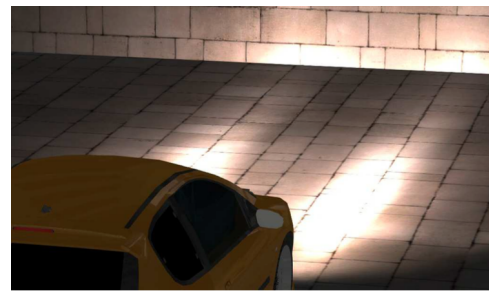
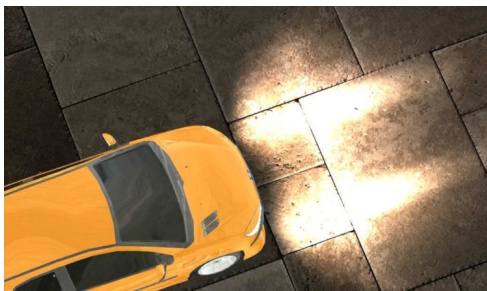


Figure 9. Our new light importance sampling technique estimates direct lighting interactively (7-9 fps) with only 200 samples per pixel that are distributed among the different images of the light field luminaire. The car headlights are represented by the same light field composed of 11×9 images (256×256 pixels).

7.1.5. Exact Relations between Wave and Ray Aberrations

The aberrations of an optical system can be described in terms of the wave aberrations, defined as the departure from the ideal spherical wavefront; or the ray aberrations, which are in turn the deviations from the paraxial ray intersections measured in the image plane. The classical connection between the two descriptions is an approximation, the error of which has, so far, not been quantified analytically.

We derive [13] exact analytical equations for computing the wavefront surface, the aberrated ray directions, and the transverse ray aberrations in terms of the wave aberrations (a.k.a., Optical Path Difference) and the reference sphere. We introduce precise conditions for a function to be an OPD function, show that every such function has an associated wavefront, and study the error arising from the classical approximation. We establish strict conditions for the error to be small. We illustrate our results with numerical simulations. Our results show that large numerical apertures and OPD functions with strong gradients yield larger approximation errors.

7.2. From Acquisition to Display

7.2.1. Lytro Microscope

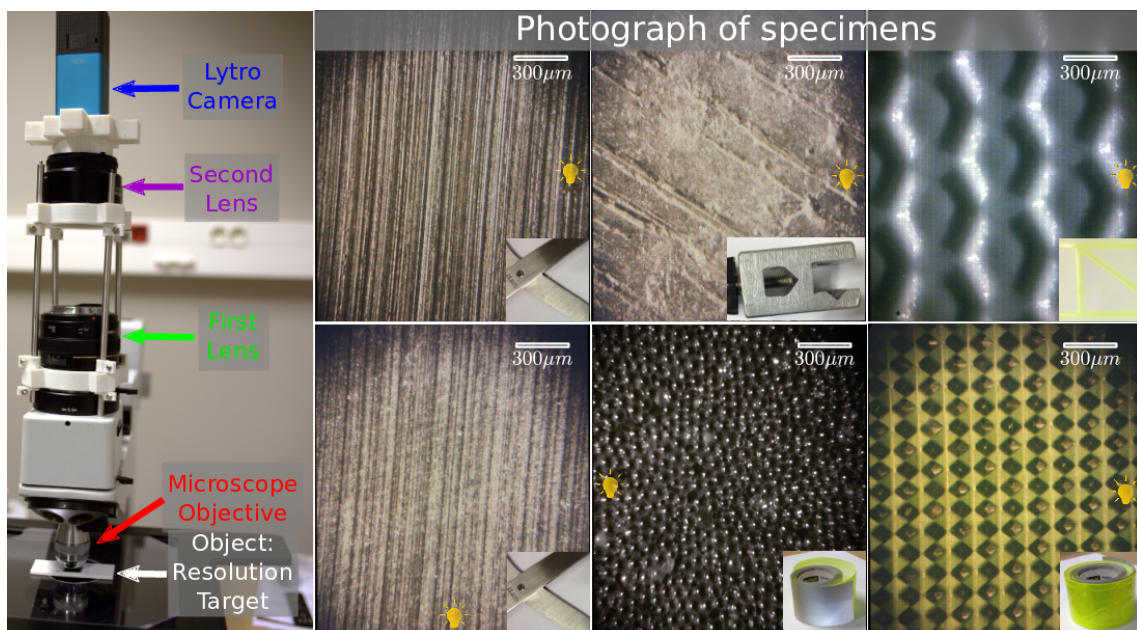


Figure 10. Light field microscopy with a consumer light field camera. Light fields can be beneficial for microscopic applications because they provide 3D information on a sample. Access to the technology, has, however, been limited by the need for custom-building the device. Our work enables an easy entry-level experimentation with the technology. (left) Light field microscope with a consumer light field camera. (right) Example specimens photographed with our system.

We explore [22] the use of inexpensive consumer light-field camera technology for the purpose of light-field microscopy. Our experiments are based on the Lytro (first generation) camera. Unfortunately, the optical systems of the Lytro and those of microscopes are not compatible, leading to a loss of light-field information due to angular and spatial vignetting when directly recording microscopic pictures. We therefore consider an

adaptation of the Lytro optical system. We demonstrate that using the Lytro directly as an ocular replacement, leads to unacceptable spatial vignetting. However, we also found a setting that allows the use of the Lytro camera in a virtual imaging mode which prevents the information loss to a large extent. We analyze the new virtual imaging mode and use it in two different setups for implementing light-field microscopy using a Lytro camera. As a practical result, we show that the camera can be used for low magnification work, as e.g. common in quality control, surface characterization, etc. (cf. Figure 10) We achieve a maximum spatial resolution of about 6.25 micrometers, albeit at a limited SNR for the side views.

7.3. Editing and Modeling

7.3.1. MatCap Decomposition for Dynamic Appearance Manipulation

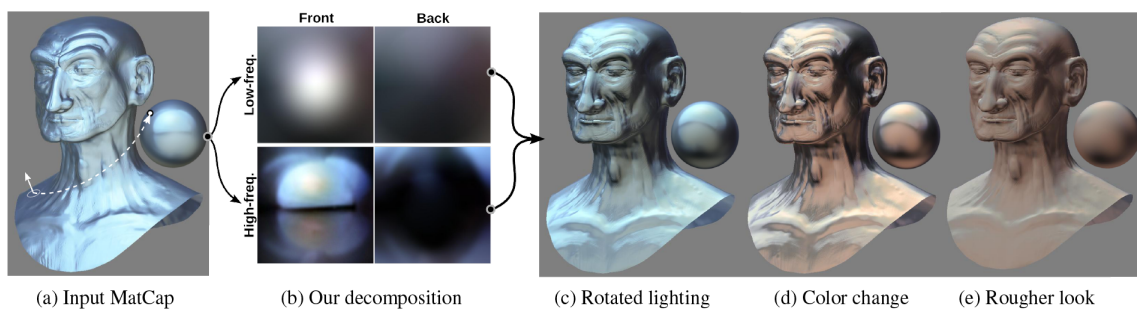


Figure 11. Our method decomposes a MatCap into a representation that permits dynamic appearance manipulation via image filters and transforms. (a) An input MatCap applied to a sculpted head model (with a lookup based on screen-space normals). (b) The low- & high-frequency (akin to diffuse & specular) components of our representation stored in dual paraboloid maps. (c) A rotation of our representation orients lighting toward the top-left direction. (d) Color changes applied to each component. (e) A rougher-looking material obtained by blurring, warping and decreasing the intensity of the high-frequency component.

In sculpting software, MatCaps are often used by artists as a simple and efficient way to design appearance. Similar to LitSpheres, they convey material appearance into a single image of a sphere, which can be easily transferred to an individual 3D object. Their main purpose is to capture plausible material appearance without having to specify lighting and material separately. However, this also restricts their usability, since material or lighting cannot later be modified independently. Manipulations as simple as rotating lighting with respect to the view are not possible. In [23], we show how to decompose a MatCap into a new representation that permits dynamic appearance manipulation. We consider that the material of the depicted sphere acts as a filter in the image, and we introduce an algorithm that estimates a few relevant filter parameters interactively. We show that these parameters are sufficient to convert the input MatCap into our new representation, which enables real-time appearance manipulations through simple image re-filtering operations. This includes lighting rotations, the painting of additional reflections, material variations, selective color changes and silhouette effects that mimic Fresnel or asperity scattering (cf. Figure 11).

7.3.2. Dynamic On-Mesh Procedural Generation

In collaboration with Technicolor, we developed a method to generate procedural models with global structures, such as growth plants, on existing surfaces at interactive time [18]. Our approach extends shape grammars to enable context-sensitive procedural generation on the GPU. To this end, we unified the representation of external contexts as texture maps, which can be spatially varying parameters controlling the grammar expansion through very fast texture fetches (e.g., a density map). External contexts also include the shape of the



Figure 12. Given a base mesh and a procedural grammar of ivy growth, our GPU-based marching rule generated the ivy geometry on-the-fly in parallel with interactive performance. In this example, the grammar expansion is guided through a user-friendly painting interface.

underlying surface itself that we represent as a texture atlas of geometry images. Extrusion along the surface is then performed by a marching rule working in texture space using indirection pointers. We also introduce a lightweight deformation mechanism of the generated geometry maintaining a C1 continuity between the terminal primitives while taking into account the shape and trajectory variations. Our method is entirely implemented on the GPU and it allows to dynamically generate highly detailed models on surfaces at interactive time (cf. Figure 12). Finally, by combining marching rules and generic contexts, users can easily guide the growing process by directly painting on the surface with a live feedback of the generated model. This provides friendly editing in production environments.

7.3.3. Boolean on general 3D meshes

Computing Boolean operations (Booleans) of 3D polyhedra/meshes is a basic and essential task in many domains, such as computational geometry, computer-aided design, and constructive solid geometry. Booleans are challenging to compute when dealing with meshes, because of topological changes, geometric degeneracies, etc. Most prior art techniques either suffer from robustness issues, deal with a restricted class of input/output meshes, or provide only approximate results.

We overcome these limitations and introduced an exact and robust approach performing on general surface meshes (closed and orientable) [11]. Our method is based on a few geometric and topological predicates that allow to handle all input/output cases considered as degenerate in existing solutions, such as voids, non-manifold, disconnected, and unbounded meshes, and to robustly deal with special input configurations. Our experimentation showed that our more general approach is also more robust and more efficient than Maya's implementation ($\times 3$), CGAL's robust Nef polyhedra ($\times 5$), and recent plane-based approaches.

During this work, we also developed a complete benchmark intended to validate Boolean algorithms under relevant and challenging scenarios, and we successfully ascertain both our algorithm and implementation with it.

7.3.4. Extending MLS surfaces

Moving least squares (MLS) surface approximation is a popular tool for the processing and reconstruction of non-structured and noisy point clouds. We introduce [14] a new variant improving the approximation

quality when the underlying surface is assumed to be locally developable, which is often the case in point clouds coming from the acquisition of manufactured objects. Our approach follows Levin's classical MLS procedure: the point cloud is locally approximated by a bivariate quadratic polynomial height-field defined in a local tangent frame. The a priori developability knowledge is introduced by constraining the fitted polynomials to have a zero-Gaussian curvature leading to the actual fit of so-called parabolic cylinders. When the local developability assumption cannot be made unambiguously, our fitted parabolic cylinders seamlessly degenerate to linear approximations. We show that our novel MLS kernel reconstructs more locally-developable surfaces than previous MLS methods while being faithful to the data.

MAVERICK Project-Team

6. New Results

6.1. Single Scattering in participating media with refractive boundaries

Participant: Nicolas Holzschuch [contact].

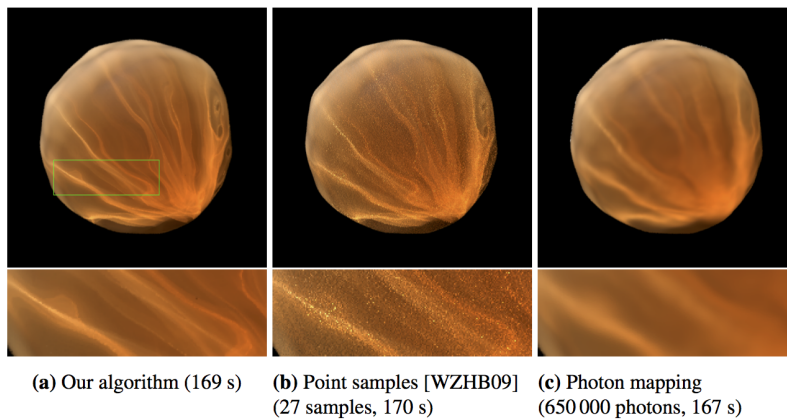


Figure 2. Single scattering: comparison between our algorithm and existing methods (equal computation time) on a translucent sphere illuminated by a point light source from behind.

Volume caustics are high-frequency effects appearing in participating media with low opacity, when refractive interfaces are focusing the light rays (see Figure 2). Refractions make them hard to compute, since screen locality does not correlate with spatial locality in the medium. We have developed a new method for accurate computation of single scattering effects in a participating media enclosed by refractive interfaces. Our algorithm is based on the observation that although radiance along each camera ray is irregular, contributions from individual triangles are smooth. Our method gives more accurate results than existing methods, faster. It uses minimal information and requires no precomputation or additional data structures. This paper was published in the *Computer Graphics Forum* journal [3] and presented at the *Eurographics Symposium on Rendering*.

6.2. Diffraction effects in reflectance properties

Participant: Nicolas Holzschuch [contact].

Reflectance properties express how objects in a virtual scene interact with light. They control the appearance of the object: whether it looks shiny or not, it has a metallic or plastic appearance. The reflectance model (BRDF) is essential for photorealistic pictures. Measured reflectance provide high realism, at the expense of memory cost. Parametric models are compact, but it is difficult to find the right parameters from measured reflectance.

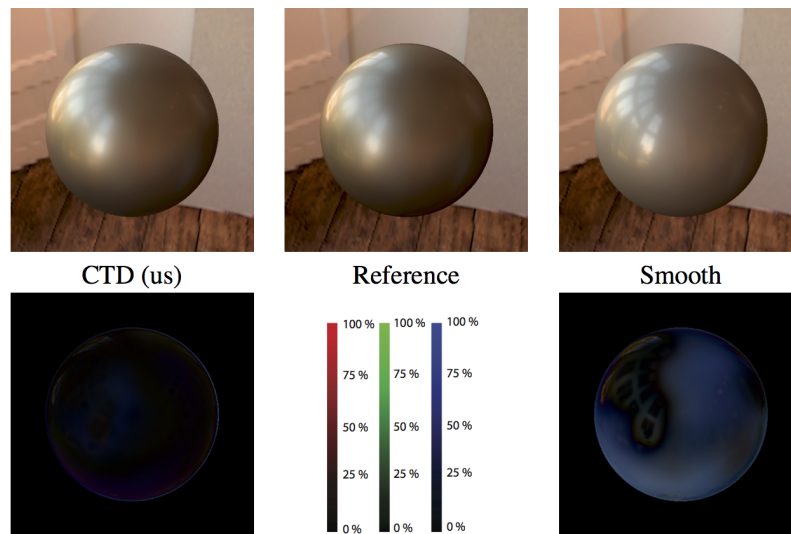


Figure 3. Comparison between our model combining reflection and diffraction (left), measured data (center) and state-of-the-art model (right).

Many parametric models are based on a physical representation of the surface micro-geometry and how it interacts with incoming light. The Cook-Torrance model assumes that light follows the principles of optical geometry: it is reflected by the surface micro-geometry but also potentially occluded. The diffraction model assumes that the micro-geometry diffracts the incoming light. This reflectance model has an intrinsic wavelength dependency. Previous experiments have shown that fitting measured materials to parametric models is hard. Heuristic models based on either Cook-Torrance or diffraction are complex, with many parameters. Our research has shown that both effects (optical geometry and diffraction) are present in most measured materials [6]. Based on this knowledge, we have proposed a new reflectance model, that accurately represents measured reflectance [10]. This model combines optical geometry for the specular peak and diffraction effects for wide-angle scattering.

6.3. Efficient and Accurate Spherical Kernel Integrals using Isotropic Decomposition

Participant: Cyril Soler [contact].

Spherical filtering is fundamental to many problems in image synthesis, such as computing the reflected light over a surface or anti-aliasing mirror reflections over a pixel. This operation is challenging since the profile of spherical filters (e.g., the view-evaluated BRDF or the geometry-warped pixel footprint, above) typically exhibits both spatial-and rotational-variation at each pixel, precluding precomputed solutions. We accelerate complex spherical filtering tasks using isotropic spherical decomposition (ISD), decomposing spherical filters into a linear combination of simpler isotropic kernels. Our general ISD is flexible to the choice of the isotropic kernels, and we demonstrate practical realizations of ISD on several problems in rendering: shading and prefiltering with spatially-varying BRDFs, anti-aliasing environment mapped mirror reflections, and filtering of noisy reflectance data. Compared to previous basis-space rendering solutions, our shading solution generates ground truth-quality results at interactive rates, avoiding costly reconstruction and large approximation errors. This paper was published in *ACM Transactions on Graphics* [4] and presented at Siggraph Asia 2015.

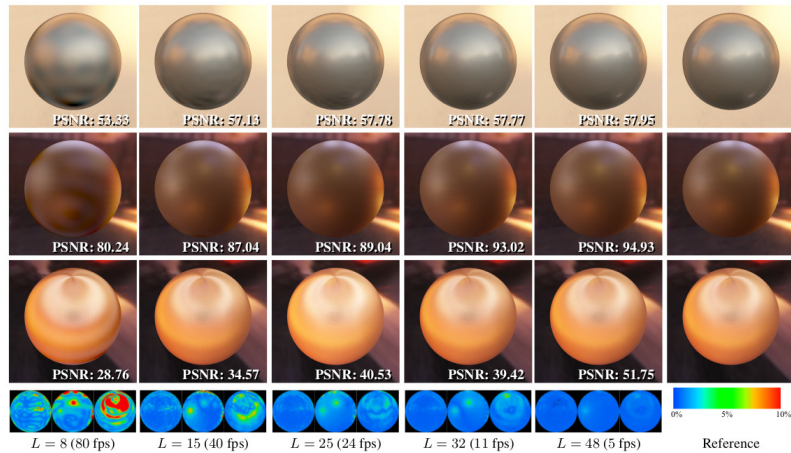


Figure 4. Convergence of spherical ISD shading for increasing L . Top to bottom: isotropic alum-bronze (with *pisa* illumination), isotropic gold-paint and anisotropic yellow-satin (both using *grace cathedral* illumination).

Reference images were ray-traced using 300K samples per pixel. False color error images in the bottom row visually illustrate the convergence of our RZH approximation to the reference rendering. Note that the dark region in the center of the spheres on the last row of renderings is indeed part of the underlying reflectance input data.

6.4. Color transfer guided by summary statistics

Participants: Benoit Arbelot, Thomas Hurtut, Romain Vergne [contact], Joëlle Thollot.



Color transfer

Colorization

Figure 5. Our framework allows for automatic local color transfer (left) and colorization (right) based on textural properties.

We have targeted two related color manipulation problems: *Color transfer* for modifying an image colors and *colorization* for adding colors to a greyscale image. Automatic methods for these two applications propose to modify the input image using a reference that contains the desired colors. Previous approaches usually do not target both applications and suffer from two main limitations: possible misleading associations between input and reference regions and poor spatial coherence around image structures. In this paper, we propose a unified framework that uses the textural content of the images to guide the color transfer and colorization. Our method introduces an edge-aware texture descriptor based on region covariance, allowing for local color

transformations. We show that our approach is able to produce results comparable or better than state-of-the-art methods in both applications. This work was presented at the AFIG conference [7]. An extended version is available as a research report [9].

6.5. Programmable 2D Arrangements for Element Texture Design

Participants: Hugo Loi, Thomas Hurtut, Romain Vergne, Joëlle Thollot [contact].

We introduce a programmable method for designing stationary 2D arrangements for element textures, namely textures made of small geometric elements. These textures are ubiquitous in numerous applications of computer-aided illustration. Previous methods, whether they be example-based or layout-based, lack control and can produce a limited range of possible arrangements. Our approach targets technical artists who will design an arrangement by writing a script. These scripts are using three types of operators: partitioning operators for defining the broad-scale organization of the arrangement, mapping operators for controlling the local organization of elements, and merging operators for mixing different arrangements. These operators are designed so as to guarantee a stationary result meaning that the produced arrangements will always be repetitive. We show (see Figure 10) that this simple set of operators is sufficient to reach a much broader variety of arrangements than previous methods. Editing the script leads to predictable changes in the synthesized arrangement, which allows an easy iterative design of complex structures. Finally, our operator set is extensible and can be adapted to application-dependent needs. This work is available as a research report [11]

6.6. Piecewise polynomial Reconstruction of Scalar Fields from Simplified Morse-Smale Complexes

Participants: Léo Allemand-Giorgis, Georges-Pierre Bonneau [contact].

Morse-Smale (MS) complexes have been proposed to visualize topological features of scalar fields defined on manifold domains. Herein, three main problems have been addressed in the past: (a) efficient computation of the initial combinatorial structure connecting the critical points; (b) simplification of these combinatorial structures; (c) reconstruction of a scalar field in accordance to the simplified Morse-Smale complex. The present paper faces the third problem by proposing a novel approach for computing a scalar field coherent with a given simplified MS complex that privileges the use of piecewise polynomial functions. Based on techniques borrowed from shape preserving design in Computer Aided Geometric Design, our method constructs the surface cell by cell using piecewise polynomial curves and surfaces. We present the benefit and limitations of using polynomials for reconstruction surfaces from topological data. This research was published in a book chapter [8].

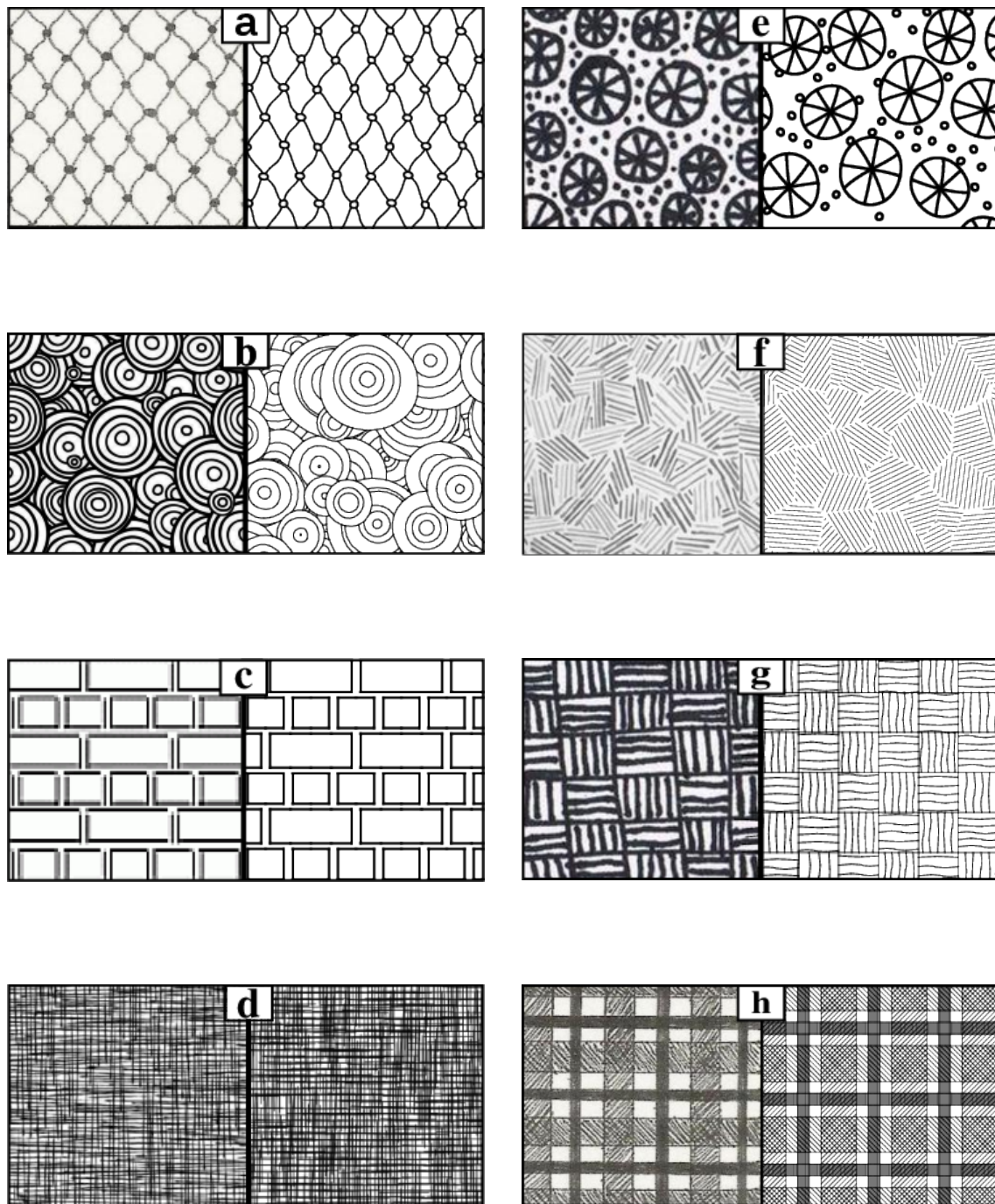


Figure 6. **Element textures commonly used.** These textures can be found in professional art (d,g,h), casual art (a,e,f), technical productions such as Computer-Assisted Design illustration tools (c), and textile industry (b). For each example, we show a hand-drawn image (left), and our synthesized reproduction of its geometric arrangement (right). (a,b,c) Classic regular distributions with contact, overlap and no adjacency between elements respectively. (d) Overlap of two textures creating cross hatching. (e) Non overlapping combination of two textures. (f,g,h) Complex element textures with clusters of elements. — Image credit: (d,g,h) “Rendering in Pen and Ink: The Classic Book On Pen and Ink Techniques for Artists, Illustrators, Architects, and Designers” ; (a,e) Profusion Art [profusionart.blogspot.com] ; (f) Hayes’ Art Classes [hayesartclasses.blogspot.com] ; (c) CompugraphX [www.compugraphx.com] ; (b) 123Stitch [www.123stitch.com].

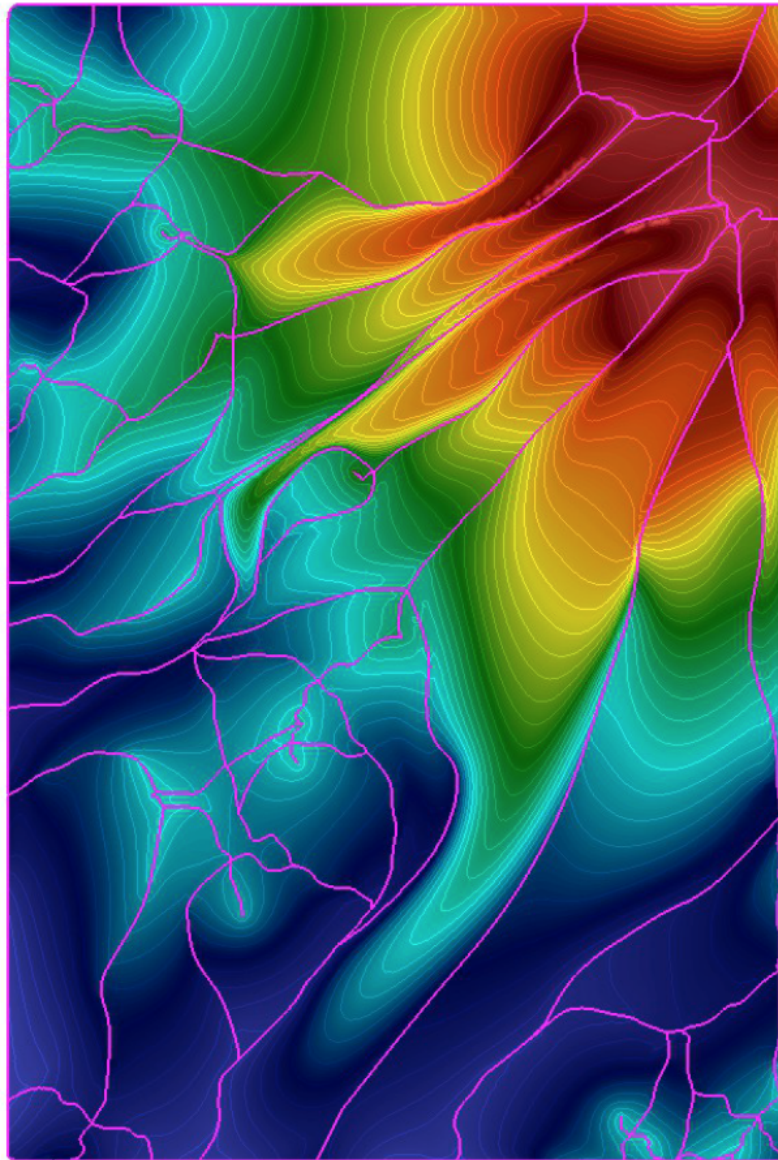


Figure 7. A function is reconstructed from its Morse-Smale complex (in purple). Inside the cells the function is monotonic so that no critical points are inserted, as can be seen from the isocontours in white. This technique is useful in Visualization whenever critical points in the data are important.

MIMETIC Project-Team

7. New Results

7.1. Biomechanics for motion analysis-synthesis

Participants: Charles Pontonnier, Georges Dumont, Steve Tonneau, Franck Multon, Julien Pettré, Ana Lucia Cruz Ruiz, Antoine Muller.

Ana-Lucia Cruz-Ruiz has been recruited as a PhD student since november 2013. The goal of this thesis is to define and evaluate muscle-based controllers for avatar animation. We developed an original control approach to reduce the redundancy of the musculoskeletal system for motion synthesis, based on the muscle synergy theory. For this purpose we ran an experimental campaign of overhead throwing motions. We recorded the muscle activity of 10 muscles of the arm and the motion of the subjects. Thanks to a synergy extraction algorithm, we extracted a reduced set of activation signals corresponding to the so called muscle synergies and used them as an input in a forward dynamics pipeline. Thanks to a two stage optimization method, we adapted the model's muscle parameters and the synergy signals to be as close as possible of the recorded motion. The results are compelling and ask for further developments [9], [24].

We are also developing an analysis pipeline thanks to the work of Antoine Muller. This pipeline aims at using a modular and multiscale description of the human body to let users be able to analyse human motion. For now, the pipeline is able to assemble different biomechanical models in a convenient descriptive graph [15], Calibrate those models thanks to experimental data [30] and run inverse dynamics to get joint torques from experimental motion capture data [14].

7.2. VR and Ergonomics

Participants: Charles Pontonnier, Georges Dumont, Pierre Plantard, Franck Multon.

The use of virtual reality tools for ergonomics applications is a very important challenge in order to generalize the use of such devices for the design of workstations.

We deeply assessed the propensity of a virtual reality immersive room and classical interaction devices to evaluate properly the physical risk factors associated to assembly tasks. For this purpose, we compared tasks realized in real and virtual environment in terms of shoulder kinematics and muscular activity [20] and in terms of controlled kinematical variables, on the basis of the uncontrolled manifold theory [31]. Results show that there is less difference between real and virtual conditions than between individuals, that make us think that such a virtual environment can be used to assess this type of task.

7.3. Interactions between walkers

Participants: Anne-Hélène Olivier, Armel Créteil, Julien Bruneau, Richard Kulpa, Sean Lynch, Julien Pettré.

Interaction between people, and especially local interaction between walkers, is a main research topic of MIMETIC. We propose experimental approaches using both real and virtual environments. This year, we developed new experiments in our immersive platform. First, we investigated obstacle avoidance behavior during real walking in a large immersive projection setup [22]. We analyze the walking behavior of users when avoiding real and virtual static obstacles. Indeed, CAVE-like immersive projection environments enable users to see both virtual and real objects, including the user's own body. With recent advances in VR technologies it becomes possible to build large-scale tracked immersive projection environments, which enable users to control their position in a large region of interest by real walking. In such environments virtual and real objects as well as multiple users or avatars may coexist in the same interaction space. Hence, it becomes important to gain an understanding of how the user's behavior is affected by the differences in perception and affordances of such real and virtual obstacles. We consider both anthropomorphic and inanimate objects,

each having his virtual and real counterpart. The results showed that users exhibit different locomotion behaviors in the presence of real and virtual obstacles, and in the presence of anthropomorphic and inanimate objects. Precisely, the results showed a decrease of walking speed as well as an increase of the clearance distance (i. e., the minimal distance between the walker and the obstacle) when facing virtual obstacles compared to real ones. Moreover, users act differently due to their perception of the obstacle: users keep more distance when the obstacle is anthropomorphic compared to an inanimate object and when the orientation of anthropomorphic obstacle is from the profile compared to a front position. However, although we observed differences in collision avoidance behavior between real and virtual obstacles, which indicate biases of natural locomotion introduced by the setup, their magnitude seem lower compared to typical results found in HMD environments. This suggests that although the user's behavior in mixed environments varies depending on the nature of the stimulus, the user's locomotion behavior and the management of his/her interaction space is comparable with the ones in real life. Considering these findings, our results open promising vistas for using large CAVE-like setups for socio-physical experiments, in particular in the fields of locomotion and behavioral dynamics.

Second, we studied interactions between an individual and a crowd [7]. When avoiding a group, a walker has two possibilities: either he goes through it or around it. Going through very dense group or around huge one would not seem natural and could break any sense of presence in a virtual environment. The aim of this work was to enable crowd simulators to correctly handle such situations. To this end, we need understanding how real humans decide to go through or around groups. As a first hypothesis, we apply the Principle of Minimum Energy (PME) on different group sizes and density. According to it, a walker should go around small and dense groups while he should go through large and sparse groups. We quantified decision thresholds. However, PME left some inconclusive situations for which the two solutions paths have similar energetic cost. In a second part, we proposed an experiment to corroborate PME decisions thresholds with real observations. We proposed using Virtual Reality to enable accurately controlling experimental factors. We considered as well the role of secondary factors in inconclusive situations. We showed the influence of the group appearance and direction of relative motion in the decision process. Finally, we draw some guidelines to integrate our conclusions to existing crowd simulators and demonstrate that spectators can perceive some improvement in the crowd animation.

This year, we also developed new experiments in real conditions by considering the interaction between a walker and a moving robot. This worked was performed in collaboration with Philippe Souères and Christian Vassallo (LAAS, Toulouse). The development of Robotics accelerated these recent years, it is clear that robots and humans will share the same environment in a near future. In this context, understanding local interactions between humans and robots during locomotion tasks is important to steer robots among humans in a safe manner. Our work is a first step in this direction. Our goal is to describe how, during locomotion, humans avoid collision with a moving robot. We study collision avoidance between participants and a non-reactive robot (we wanted to avoid the effect of a complex loop by a robot reacting to participants' motion). Our objective is to determine whether the main characteristics of such interaction preserve the ones previously observed: accurate estimation of collision risk, anticipated and efficient adaptations. We observed that collision avoidance between a human and a robot has similarities with human-human interactions (estimation of collision risk, anticipation) but also leads to major differences. Humans preferentially give way to the robot, even if this choice is not optimal with regard to motion adaptation to avoid the collision. We proposed to interpret this behavior based on the notion of perceived danger and safety. Given the difficulty to understand how a robot behaves, and the lack of experience of interactions with the robot, humans apply a conservative avoidance strategy and prefer giving way to the robot. However, it is important to note that human participants succeed in perceiving the motion of the robot (anticipation was observed, no aberrant reaction occurred). One main conclusion is that, if we control robots to move like humans, we have a risk facing unexpected situations where robot compensates and cancels humans adaptations to the robot. A robot programmed to be cooperative could be perceived as hostile. The conclusion of this study opens paths for future research. A first direction is to better understand the possible effect of this notion of danger during interactions. We believe that this notion is of even higher importance when studying interactions with vehicles: a risk of collision with a fast vehicle obviously raises higher danger. A second direction is about the design of safe robots moving among human walkers. How

the robot should adapt to others? Should it be collaborative with the risk of compensating human avoidance strategies? Should it be passive? We believe that robots should first be equipped with the ability to early detect humans avoidance strategy and adapt to it. In the near future, we want to continue our study of interactions between a robot and a human. In a first step, we plan to equip the robot with collision avoidance system which imitates real human strategies, and investigate how participants adapt to this new situation in comparison with a passive robot.

Finally, Sean Dean Lynch has been recruited as a PhD student since september 2015. This thesis concerns the visual perception of human motion during interactions in locomotor tasks. From the visual perception of someone's motion, we are able to predict the future course of this motion, interpret and anticipate his/her intentions and adapt our own motion to allow interactions. The main objective of the thesis is to identify the underlying perceptual mechanisms, i.e., the human motion cues which are necessary for an accurate understanding of others' intentions. It would allow to make significant progress in the understanding of human social behaviors. To reach these objectives, the thesis will be based on an experimental approach in virtual reality.

7.4. Motion Sensing

Participants: Franck Multon, Pierre Plantard.

Recording human activity is a key point of many applications and fundamental works. Numerous sensors and systems have been proposed to measure positions, angles or accelerations of the user's body parts. Whatever the system is, one of the main is to be able to automatically recognize and analyze the user's performance according to poor and noisy signals. Hence, recognizing and measuring human performance are important scientific challenges especially when using low-cost and noisy motion capture systems. MimeTIC has addressed the above problems in two main application domains.

Firstly, in ergonomics, we explored the use of low-cost motion capture systems, a Microsoft Kinect, to measure the 3D pose of a subject in natural environments, such as on a workstation, with many occlusions and inappropriate sensor placements. Predicting the potential accuracy of the measurement for such complex 3D poses and sensor placements is challenging with classical experimental setups. To tackle this problem, we propose [16] a new evaluation method based on a virtual mannequin. Thanks to this evaluation method, more than 500,000 configurations have been automatically tested, which is almost impossible to evaluate with classical protocols. The results show that the kinematic information obtained by the Kinect system is generally accurate enough to fill-in ergonomic assessment grids. However inaccuracy strongly increases for some specific poses and sensor positions. Using this evaluation method enabled us to report configurations that could lead to these high inaccuracies. Results obtained with the virtual mannequin are in accordance with those obtained with a real subject for a limited set of poses and sensor configuration. This knowledge can help to anticipate potential problems using a Kinect in given scenarios, and to propose methods to tackle these expected problems.

Secondly, in clinical gait analysis, we proposed a method to overcome the main limitations imposed by the low accuracy of the Kinect measurements in real medical exams. Indeed, inaccuracies in the 3D depth images leads to badly reconstructed poses and inaccurate gait event detection. In the latter case, confusion between the foot and the ground leads to inaccuracies in the foot-strike and toe-off event detection, which are essential information to get in a clinical exam. To tackle this problem we assumed that heel strike events could be indirectly estimated by searching for the extreme values of the distance between the knee joints along the walking longitudinal axis [5]. As Kinect sensor may not accurately locate the knee joint, we used anthropometrical data to select a body point located at a constant height where the knee should be in the reference posture. Compared to previous works using a Kinect, heel strike events and gait cycles are more accurately estimated, which could improve global clinical gait analysis frameworks with such a sensor. Once these events are correctly detected, it is possible to define indexes that enables the clinician to have a rapid state of the quality of the gait. We proposed [4] a new method to asses gait asymmetry based on depth images, to decrease the impact of errors in the Kinect joint tracking system. It is based on the longitudinal

spatial difference between lower-limb movements during the gait cycle. The movement of artificially impaired gaits was recorded using both a Kinect placed in front of the subject and a motion capture system. The proposed longitudinal index distinguished asymmetrical gait ($p < 0.001$), while other symmetry indices based on spatiotemporal gait parameters failed using such Kinect skeleton measurements. This gait asymmetry index measured with a Kinect is low cost, easy to use and is a promising development for clinical gait analysis.

7.5. Virtual Human Animation

Participants: Julien Pettré, Franck Multon, Steve Tonneau.

Multipled locomotion in cluttered environments is addressed as the problem of planning acyclic sequences of contacts, that characterize the motion. In order

to overcome the inherent combinatorial difficulty of the problem, we separate it in two subproblems [34]: first, planning a guide trajectory for the root of the robot and then, generating relevant contacts along this trajectory. This paper proposes theoretical contributions to these two subproblems. We propose a theoretical characterization of the guide trajectory, named “true feasibility”, which guarantee that a guide can be mapped into the contact manifold of the robot. As opposed to previous approaches, this property makes it possible to assert the relevance of a guide trajectory without explicitly computing contact configurations, as proposed in our previous works. This property can be efficiently checked by a sample-based planner (e.g. we implemented a visibility PRM). Since the guide trajectories that we characterized are easily mapped to a valid sequence of contacts, we then focused on how to select a particular sequence with desirable properties, such as robustness, efficiency and naturalness, only considered for cyclic locomotion so far. Based on these novel theoretical developments, we implemented a complete acyclic contact planner and demonstrate its efficiency by producing a large variety of movements with three very different robots (humanoid, insectoid, dexterous hand) in five challenging scenarios. The planner is very efficient in quality of the produced movements and in computation time: given a computed RB-PRM, a legged figure or a dexterous hand can generate its motion in real time. This result outperforms any previous acyclic contact planner.

7.6. VR and sports

Participants: Richard Kulpa, Benoit Bideau, Franck Multon, Anne-Hélène Olivier.

Athletes’ performances are influenced by internal and external factors, including their psychological state and environmental factors, especially during competition. As a consequence, current training programs include stress management. In this work [3], we explore whether highly immersive systems can be used for such training programs. First, we propose methodological guidelines to design sport training scenarios both on considering the elements that a training routine must have and how external factors might influence the participant. The proposed guidelines are based on Flow and social-evaluative threat theories. Second, to illustrate and validate our methodology, we designed an experimental setup reproducing a 10 m Olympic pistol shooting. We analyzed whether changes in the environment are able to induce changes in user performance, physiological responses, and the subjective perception of the task. The simulation included stressors in order to raise a social-evaluative threat, such as aggressive public behavior or unforced errors, increasing the pressure while performing the task. The results showed significant differences in their subjective impressions, trends in the behavioral and physiological data were also observed. Taken together, our results suggest that highly immersive systems could be further used for training in sports.

Among the stimuli, visual information uptake is a fundamental element of sports involving interceptive tasks. Several methodologies, like video and methods based on virtual environments, are currently employed to analyze visual perception during sport situations. Both techniques have advantages and drawbacks. We made an experiment to determine which of these technologies may be preferentially used to analyze visual information uptake during a sport situation [21]. To this aim, we compared a handball goalkeeper’s performance using two standardized methodologies: video clip and virtual environment. We examined this performance for two response tasks: an uncoupled task (goalkeepers show where the ball ends) and a coupled task (goalkeepers

try to intercept the virtual ball). Variables investigated in this study were percentage of correct zones, percentage of correct responses, radial error and response time. The results showed that handball goalkeepers were more effective, more accurate and started to intercept earlier when facing a virtual handball thrower than when facing the video clip. These findings suggested that the analysis of visual information uptake for handball goalkeepers was better performed by using a 'virtual reality'-based methodology.

In a previous work, we analyzed the performance of beginners as they shot basketball free throws using various immersive conditions. Our results supported the assumption that natural complex motor behavior is possible in a VE, with little motor adaptation. The ultimate goal of our work is to design a VE training system for basketball free throws, so in this article we compare the performance of beginners making free throws in various visual conditions (first- versus third-person views using a large-screen immersive display) with that of expert players in the real world [8]. The key idea is to analyze how different visual conditions affect the performance of novices and to what extent it enables them to match the experts' performance.

Distance underestimation or any other perceptual disturbance in VR makes people adapt to the task at hand. The users in our study reached the same success rate by finding a new way to throw the ball, despite this incongruity between perception and action. The main observations reported in this article reinforce the conclusions in previous work, stating that 3PP is more efficient for certain tasks, but further work is required to test this result against other types of training conditions.

Finally, we worked on a transportable virtual reality system to analyse sports situations [6]. We proposed an original methodology to study the action of a goalkeeper facing a free kick. This methodology is based on a virtual reality setup in which a real goalkeeper is facing a virtual player and a virtual defensive wall. The setup has been improved to provide a total freedom of movement to the goalkeeper in order to have a realistic interaction between the goalkeeper and the player. The goalkeeper's movements are captured in real-time to accurately analyze his reactions. Such a methodology not only represents a valuable research tool but also provides a relevant training tool. Using this setup, this paper shows that goalkeepers are more performant during free kick with a wall composed of 5 defenders whatever its position.

7.7. Scheduling activities under spatial and temporal constraints

Participants: Fabrice Lamarche, Carl-Johan Jorgensen.

This work focusses on generating statistically consistent behaviors that can be used to pilot crowd simulation models over long periods of time, up to multiple days [1]. In real crowds, people's behaviors mainly depend on the activities they intend to perform. The way this activity is scheduled rely on the close interaction between the environment, space and time constraints associated with the activity and personal characteristics of individuals. Compared to the state of the art, our model better handle this interaction.

Our main contributions lie in the cdomain of activity scheduling and path planning. First, we proposed an individual activity scheduling process and its extension to cooperative activity scheduling. Based on descriptions of the environment, of intended activities and of agents' characteristics, these processes generate a task schedule for each agent. Locations where the tasks should be performed are selected and a relaxed agenda is produced. This task schedule is compatible with spatial and temporal constraints associated with the environment and with the intended activity of the agent and of other cooperating agents. It also takes into account the agents personal characteristics, inducing diversity in produced schedules. We showed that this model produces schedules statistically coherent with the ones produced by humans in the same situations. Second, we proposed a hierarchical path-planning process. It relies on an automatic environment analysis process that produces a semantically coherent hierarchical representation of virtual cities. The hierarchical nature of this representation is used to model different levels of decision making related to path planning. A coarse path is first computed, then refined during navigation when relevant information is available. It enable the agent to seamlessly adapt its path to unexpected events. Finally, those models have been included in a simulation platform that is able to simulate several thousand of pedestrians performing their daily activities in real-time. In order to deal with unexpected events, a process enabling adaptations of the pedestrian behavior have been designed. Those adaptations range from path modification to schedule adaptation according to the observed situation.

The proposed model handles long term rational decisions driving the navigation of agents in virtual cities. It considers the strong relationship between time, space and activity to produce more credible agents' behaviors. It can be used to easily populate virtual cities in which observable crowd phenomena emerge from individual activities.

7.8. Shoulder biomechanics

Participant: Armel Crétual [contact].

Shoulder hyperlaxity (SHL) is considered a main risk factor for shoulder instability and can be associated with different clinical shoulder instability presentations, such a multidirectional instability or unstable painful shoulder. Interestingly, quantification of shoulder laxity and hyperlaxity, particularly during physical examination, still remains an unsolved problem. Indeed, it is still frequently evaluated only through mono-axial amplitude, in particular using external rotation of the arm whilst at the side (ER1). We previously showed that this parameter is sensitive to inter-operator variability.

Therefore, we proposed a novel way to account for global shoulder mobility, the Shoulder Configuration Space Volume (SCSV) corresponding to the reachable volume in the configuration space of the shoulder joint [10]. In mechanics and robotics, the configuration space is the set of all reachable combination of coordinates. Considering the shoulder as the single joint between thorax and humerus instead of a combination of 4 actual joints (gleno-humeral, thoraco-humeral, scapulo-thoracic and sterno-clavicular), these coordinates are based upon the three joint angles defined by the International Society of Biomechanics (ISB) recommendations as plane of elevation orientation, elevation and axial rotation.

Then, this new index was examined through correlation to shoulder signs of hyperlaxity [19] for which we have shown a link with instability in patients who received a surgical procedure [18].

7.9. The Toric Space: a novel representation for camera control applications

Participants: Marc Christie, Christophe Lino, Quentin Galvane.

Many types of computer graphics applications such as data visualization or virtual movie production require users to position and move viewpoints in 3D scenes to effectively convey visual information or tell stories. The desired viewpoints and camera paths need to satisfy a number of visual properties (e.g. size, vantage angle, visibility, and on-screen position of targets). Yet, existing camera manipulation tools only provide limited interaction methods and automated techniques remain computationally expensive.

We introduce the *Toric space*, a novel and compact representation for intuitive and efficient virtual camera control. We first show how visual properties are expressed in this Toric space and propose an efficient interval-based search technique for automated viewpoint computation. We then derive a novel screen-space manipulation technique that provides intuitive and real-time control of visual properties. Finally, we propose an effective viewpoint interpolation technique which ensures the continuity of visual properties along the generated paths. The proposed approach (i) performs better than existing automated viewpoint computation techniques in terms of speed and precision, (ii) provides a screen-space manipulation tool that is more efficient than classical manipulators and easier to use for beginners, and (iii) enables the creation of complex camera motions such as long takes in a very short time and in a controllable way. As a result, the approach should quickly find its place in a number of applications that require interactive or automated camera control such as 3D modelers, navigation tools or games. The paper has been presented at SIGGRAPH 2015 (see [12] for more details).

We then rely on this Toric Space representation to construct optimal camera paths (optimal in the satisfaction of visual properties along the path). Indeed, when creating real or computer graphics movies, the questions of how to layout elements on the screen, together with how to move the cameras in the scene are crucial to properly conveying the events composing a narrative. Though there is a range of techniques to automatically compute camera paths in virtual environments, none have seriously considered the problem of generating realistic camera motions even for simple scenes. Among possible cinematographic devices, real cinematographers often rely on camera rails to create smooth camera motions which viewers are familiar with. Following

this practice, we have proposed a method for generating virtual camera rails and computing smooth camera motions on these rails. Our technique analyzes characters motion and user-defined framing properties to compute rough camera motions which are further refined using constrained-optimization techniques. Comparisons with recent techniques demonstrate the benefits of our approach and opens interesting perspectives in terms of creative support tools for animators and cinematographers. See [25] for more details.

TO address the more general problem of solving contradicting visual properties, novel ways of aggregating functions has also been proposed [33].

7.10. Data-driven Virtual Cinematography

Participant: Marc Christie.

Our propelling motivation here is to rely on existing data from real movies (automatically extracted or manually annotated), to propose better better and better framing techniques.

We first contributed to the problem of automated editing, by reproducing elements of cinematographic style. Automatically computing a cinematographic consistent sequence of shots over a set of actions occurring in a 3D world is a complex task which requires not only the computation of appropriate shots (viewpoints) and appropriate transitions between shots (cuts), but the ability to encode and reproduce elements of cinematographic style. Models proposed in the literature, generally based on finite state machine or idiom-based representations, provide limited functionalities to build sequences of shots. These approaches are not designed in mind to easily learn elements of cinematographic style, nor do they allow to perform significant variations in style over the same sequence of actions. We have proposed a model for automated cinematography that can compute significant variations in terms of cinematographic style, with the ability to control the duration of shots and the possibility to add specific constraints to the desired sequence. The model is parameterized in a way that facilitates the application of learning techniques. By using a Hidden Markov Model representation of the editing process, we have demonstrated the possibility of easily reproducing elements of style extracted from real movies. Results comparing our model with state-of-the-art first order Markovian representations illustrate these features, and robustness of the learning technique is demonstrated through cross-validation. See [13] for more details.

We also proposed a tool to ease the process of annotating cinematographic content, for the purposes of both film analysis, and film synthesis [29]. The work relies on the proposition of a film language that extends previous representations such as PSL (Prose Storyboard Language) by integrating the editing aspects, through the notion of cinematographic “techniques” described as patterns of shots.

The proposed language, named “Patterns”, is described in more details in [35]. Our language can express the aesthetic properties of framing and shot sequencing, and of camera techniques used by real directors. Patterns can be seen as the semantics of camera transitions from one frame to another. The language takes an editors view of on-screen aesthetic properties: the size, orientation, relative position, and movement of actors and objects across a number of shots. We have illustrated this language through a number of examples and demonstrations. Combined with camera placement algorithms, we demonstrated the language’s capacity to create complex shot sequences in data-driven generative systems for 3D storytelling applications.

7.11. Logic control in interactive storytelling

Participants: Marc Christie, Hui-Yin Wu.

With the rising popularity of engaging storytelling experiences in gaming arises the challenge of designing logic control mechanisms that can adapt to increasingly interactive, immersive, and dynamic 3D gaming environments. Currently, branching story structures are a popular choice for game narratives, but can be rigid, and authoring mistakes may result in dead ends at runtime. This calls for automated tools and algorithms for logic control over flexible story graph structures that can check and maintain authoring logic at a reduced cost while managing user interactions at runtime. In this work we introduce a graph traversal method for logic control over branching story structures which allow embedded plot lines. The mechanisms are designed to

assist the author in specifying global authorial goals, evaluating the sequence of events, and automatically managing story logic during runtime. Furthermore, we showed how our method can be easily linked to 3D interactive game environments through a simple example involving a detective story with a flashback. See [36] for more details.

7.12. Automatic Continuity Editing for 3-D Animation

Participants: Marc Christie, Quentin Galvane, Christophe Lino.

We have proposed an optimization-based approach for automatically creating movies from 3-D animation. The method nicely separates the work of the virtual cinematographer (placing cameras and lights to produce nice-looking views of the action) from the work of the virtual film editor (cutting and pasting shots from all available cameras). While previous work has mostly focused on the first problem, the second problem has never been addressed in full details. We have reviewed the main causes of editing errors and built a cost function for minimizing them. We made a plausible semi-Markov assumption, which results in a computationally efficient dynamic programming solution. We showed that our method generates movies that avoid many common errors in film editing, including jump cuts, continuity errors and non-motivated cuts. We also show that our method can generate movies with different paces. Combined with state-of-the-art cinematography, our approach therefore promises to significantly extend the expressiveness and naturalness of virtual movie-making. The work has been published at AAAI [27]. More details comparisons have been performed in [26].

MINT Project-Team

6. New Results

6.1. Physical and Perceptual Independence of Ultrasonic Vibration and Electro vibration for Friction Modulation

Eric Vezzoli, Wael Ben Messaoud, Michel Amberg, Betty Lemaire-Semail, Frédéric Giraud, Marie-Ange Bueno

Two different principles are available to modulate the user perceived roughness of a surface: electro vibration and ultrasonic vibration of a plate. The former enhances the perceived friction coefficient and the latter reduces it. In this work, we highlighted the independence of the two effects on the physical and perceptual point of view to confirm the increased range of sensation that can be supplied by the two coupled techniques. Firstly, a tribometric analysis of the induced lateral force on the finger by the two coupled effects has been achieved, then a study on the dynamic of the two effects is reported. In the end, a psychophysical experiment on the perception of the two coupled techniques confirms the approach.

6.2. Preliminary design of a multi-touch ultrasonic tactile stimulator

Sofiane Ghenna, Christophe Giraud-Audine, Frédéric Giraud, Michel Amberg, Betty Lemaire-Semail

Currently there is no solution able to provide a multitouch tactile stimulation based on friction reduction tactile devices. The main objective of this work is to achieve a control method which allows to have a differentiated tactile stimulation on two fingers simultaneously, by superimposing two vibration modes. The proof of concept has been established on a 1D beam, where the tactile stimulation could be differentiated on two selected positions. We have presented the key design rules, as well as the control method. Finally, a psychophysical evaluation has shown that users can detect the location of nodes and antinodes of vibration with an average success rate of 78%

6.3. Generalised modal analysis for closed-loop piezoelectric devices

Christophe Giraud-Audine, Frédéric Giraud, Michel Amberg, Betty Lemaire-Semail.

Stress in piezoelectric material can be controlled by imposing the electrical field. Thanks to a feedback, this electrical field can be a function of some strain related measurement so as to confer to the piezoelectric device a closed loop macroscopic behaviour. We address the modelling of such system by extending the modal decomposition methods to account for the closed loop. To do so the boundary conditions are modified to include the electrical feedback circuit, hence allowing a closed-loop modal analysis. A case study is used to illustrate the theory and to validate it. The main advantage of the method is that design issue such as coupling factor of the device and closed loop stability are simultaneously captured.

6.4. Pressure dependence of friction modulation in ultrasonic devices

Wael Ben Messaoud, Eric Vezzoli, Frédéric Giraud, Betty Lemaire-Semail

Ultrasonic vibrating devices are able to modulate the friction of a finger sliding on them. The underlying principles of the friction reduction are still unclear, and this work is carried out to investigate the influence of the ambient pressure on the friction modulation. A specific tactile stimulator has been used for this purpose and the friction between the finger sliding on the device has been recorded for an ambient pressure of 0.5 and 1 atm showing a significant difference for comparable experimental conditions. A comparison with the model proposed in literature is performed underlying that the squeeze film interaction can be present but not the only responsible of the friction modulation in this kind of devices.

Mjolnir Team

7. New Results

7.1. Introduction

The following sections summarize our main results of the year. For a complete list, see the list of publications at the end of this report.

7.2. HCI models, theories, and frameworks

Participants: Géry Casiez, Alix Goguey, Stéphane Huot.

Pointing is one of the most common and frequent action made with any interactive system whether it be a desktop computer, a mobile device or a wall-size display. Although it has been extensively studied in HCI, current pointing techniques provide no adequate way to select very small objects whose movements are fast and unpredictable, and theoretical tools –such as Fitts’ law– do not model unpredictable motion. To inform the design of appropriate selection techniques, we studied human performance (speed and accuracy) when selecting moving objects in a 2D environment with a standard mouse. We characterized selection performance as a function of the predictability of the moving targets, based on three parameters: the speed (S) of the target, the frequency (F) at which the target changes direction, and the amplitude (A) of those direction changes. Our results show that for a given speed, selection is relatively easy when A and F are both low or high, and difficult otherwise [22].

In spite of previous work showing the importance of understanding users’ strategies when performing tasks, HCI researchers comparing interaction techniques remain mainly focused on performance. This can be explained to some extent by the difficulty to characterize users’ strategies. To alleviate this problem, we introduced new metrics to quantify if an interaction technique introduces an object or command-oriented strategy, i.e. if users favor completing the actions on an object before moving to the next one, or in contrast, if they are reluctant to switch between commands [21]. Through a study comparing two novel interaction techniques with two from the literature, we showed that our metrics allow to replicate previous findings on users’ strategies concerning the latter.

To our knowledge, there are no general design and evaluation methodologies available for the development of digital musical instruments (DMI). One reason is the large diversity of design and evaluation contexts possible in musical interaction, e.g. is this evaluation done from the perspective of the DMI designer/manufacturer, the musician playing it, or the audience watching it be performed? With our collaborators of the MIDWAY associate team, we have analyzed all papers and posters published in the proceedings of the NIME conference from 2012 to 2014 [16]. For each publication that explicitly mentioned the term “evaluation”, we looked for: a) What targets and stakeholders were considered? b) What goals were set? c) What criteria were used? d) What methods were used? e) How long did the evaluation last? Results show different understandings of evaluation, with little consistency regarding the usage of the word. Surprisingly in some cases, not even basic information such as goal, criteria and methods were provided. Beyond the attempt to provide an idea of what “evaluation” means for the NIME community, we pushed the discussion towards how we could make a better use of it and what criteria should be used regarding each goal.

7.3. Transfer functions and latency

Participants: Géry Casiez, Alix Goguey, Stéphane Huot, Sylvain Malacria, Nicolas Roussel.

Our work on transfer functions mainly focused this year on *edge-scrolling*, which allows users to scroll a viewport while simultaneously dragging near or beyond its edge. Common implementations rely on rate control, mapping the distance between the pointer and the edge of the viewport to the scrolling velocity. While ubiquitous in operating systems, edge-scrolling had received little attention, though previous works suggested that rate control may be suboptimal for isotonic pointing devices (e.g. mice and touchpads) and space beyond the window's edge might be scarce, limiting scrolling control. To address these problems, we developed *Push-Edge* and *Slide-Edge* two position-based techniques that allow scrolling by “pushing” against the viewport edge [23]. A controlled experiment shows that our techniques reduce overshoots and offer performance improvements up to 13% over traditional edge-scrolling.

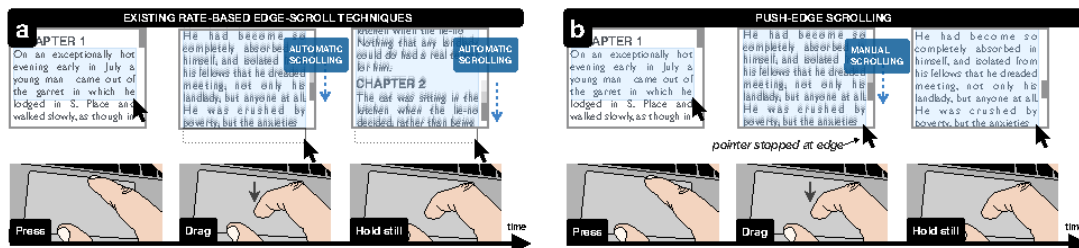


Figure 1. When selecting text with a touchpad, downward movements after crossing the viewport edge will (a) change the rate of automatic scrolling with existing techniques (rate control); or (b) manually scroll the document, stopping the pointer at the edge with push-edge scrolling (position control).

Our work on latency focused on its measurement in existing graphical user interfaces, a problem for which we developed a simple method [18], [27]. Our method consists in positioning an unmodified optical mouse on the screen while displaying and translating a particular texture to fake mouse displacements, which results in controlled mouse events. This works with most optical mice and allows accurate and real-time latency measures up to 5 times per second. The method also allows easy insertion of probes at different places in the system to investigate the sources of latency. Measurements performed on different systems, toolkits and applications notably showed that latency is affected by the operating system and system load. Substantial differences were also found between C++/GLUT and C++/Qt or Java/Swing implementations, as well as between Web browsers.

7.4. Interaction techniques

Participants: Géry Casiez, Stéphane Huot, Sylvain Malacria.

While touchpads are both widespread and expressive input devices, there has been surprisingly little research regarding how they could be used for more than simple pointer movements. In [17], we explore the design space of gesture shortcuts on touchpads and introduce four novel interaction techniques. *SpotPad* and *LociPad* rely on one-finger static gestures, but differ in their graphical representation. *ChordPad* relies on two-finger static gestures with a hierarchical representation. Finally, *InOutPad* relies on dynamic gestures crossing the edges of the touchpad. We compare the properties of these four interaction techniques and describe how they can be deployed on OS X.

The hands of virtual characters are highly complex 3D models that can be tedious and time-consuming to animate with current methods. In [14], we introduce *THING*, a novel tablet-based approach that leverages multi-touch interaction for the quick and precise control of a 3D hand's pose. The flexion/extension and abduction/adduction of the virtual fingers can be controlled individually for each finger or for several fingers in parallel through sliding motions on the tablet. We describe two variants of *THING*: *MobileTHING*, which

maps the spatial location and orientation of the tablet to that of the virtual hand, and DesktopTHING, which combines multi-touch controls of fingers with traditional mouse controls for the hand's global position and orientation. We also report on two usability studies in which we compared THING to mouse-only controls and a data glove.



Figure 2. THING enables the control of 3D hand models (in blue) by sliding fingers along sliders arranged in a morphologically-consistent pattern on the tablet's screen.

Interactive technologies have radically changed the way visual artists work, and a large portion of the artistic production has now moved from paper to the computer. However, many artists still work on paper and keep using traditional painting and drawing tools. This is not only due to resistance to progress or due to the well-known usability properties of physical tools: despite the use of pen displays, the progress of artistic stroke-rendering techniques, and the powerful and advanced functionalities of existing computer tools, they fail to fully capture the richness and variety of artistic styles supported by physical media. We interviewed four professional illustrators in their work environment. We also followed the work of an artist for a two-year period. We observed that artists mix a variety of techniques that involve specialized computer software and hardware such as Adobe Photoshop, a graphics tablet and a scanner, and traditional physical tools such as pencils, paper, and customized light tables. Our findings inspired *BricoSketch* [25], an augmented paper interface that enables illustrators to zoom into parts of their drawings and work at different levels of detail on paper. Our early results show that *BricoSketch* supports real tasks, improving productivity on paper while enhancing illustrators' creative ways of working.

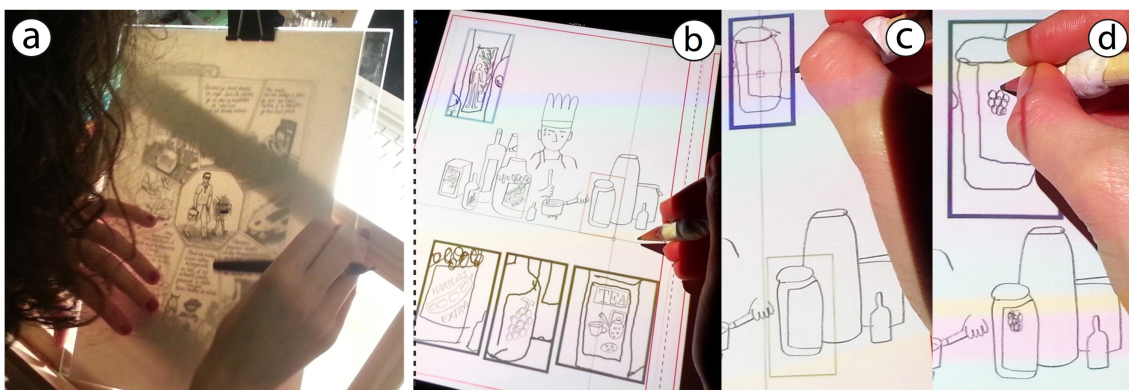


Figure 3. *BricoSketch*. (a) An artist works on layers of physical paper with a home-made light table: she draws the panels of a page for a graphic novel by using earlier sketches as guides. (b-d) The same artist uses our system to add details to different parts of her illustration through partial scaled views.

7.5. Interactive visualization

Participant: Fanny Chevalier.

The differential diagnosis of hereditary disorders is a challenging task for clinicians due to the heterogeneity of phenotypes that can be observed in patients. Existing clinical tools are often text-based and do not emphasize consistency, completeness, or granularity of phenotype reporting, which can impede clinical diagnosis and limit their utility to genetics researchers. The *PhenoBlocks* tool described in [13] is a novel visual analytics platform designed to support clinical differential diagnosis. It supports the comparison of phenotypes between patients, or between a patient and the hallmark features of a disorder. An informal evaluation with expert clinicians suggests that the visualization effectively guides the process of differential diagnosis and could reinforce the importance of complete, granular phenotypic reporting.

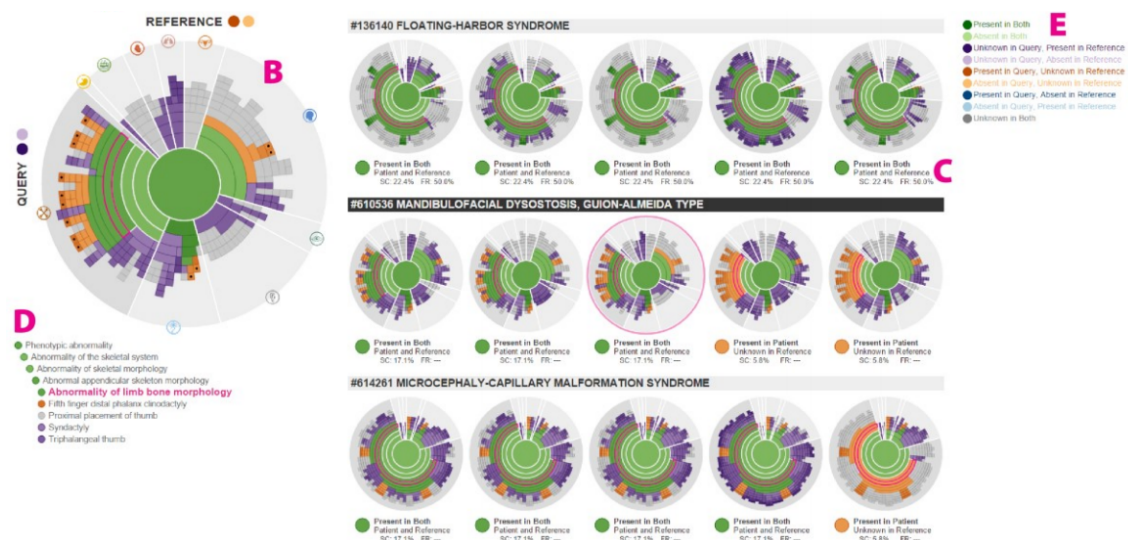


Figure 4. *PhenoBlocks* allows to compare the phenotype hierarchies of an undiagnosed query patient to a diagnosed reference patient. During differential diagnosis, clinicians use shared (green) phenotypes to gauge confidence in their diagnostic hypothesis and missing (purple) phenotypes to identify candidates for subsequent analysis.

7.6. Adaptive interfaces

Participant: Sylvain Malacria.

As news is increasingly accessed on smartphones and tablets, the need for personalizing news applications is apparent. In [19], [20], we report on a series of studies addressing key issues in the development of adaptive news interfaces. We first surveyed users' news reading preferences and behaviors. We then implemented and deployed an Android application that logs users' interactions with the application. We used the logs to train a classifier and showed that it is able to reliably recognize a user according to their reader type. Finally we evaluated alternative, adaptive user interfaces for each reader type. The evaluation demonstrates the differential benefit of the adaptation for different users and the feasibility of adaptive interfaces for news applications.

In [12], we investigate the use of a companion application on a tablet to augment viewing of information-rich television programs. The application displays a synchronized graphical abstraction of the program's content in the form of a concept map. Two experiments were conducted involving participants watching an astronomy documentary. Results show that the companion application improved participants' understanding and recall of the program. Participants were found to manage their visual attention systematically when using the companion application, and correlations were found in the way they shifted their gaze from TV screen to tablet and back in response to changes in the program content. Increasing interaction with the application disrupted understanding of the television program and visual attention. Participants were positive about the value of companion applications for understanding and recall of programs, but distraction and "knowing where to look" were significant concerns.

MORPHEO Project-Team

7. New Results

7.1. QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids

While studied over several decades, the computation of boolean operations on polyhedra is almost always addressed by focusing on the case of two polyhedra. For multiple input polyhedra and an arbitrary boolean operation to be applied, the operation is decomposed over a binary CSG tree, each node being processed separately in quasilinear time. For large trees, this is both error prone due to intermediate geometry and error accumulation, and inefficient because each node yields a specific overhead. We introduce a fundamentally new approach to polyhedral CSG evaluation, addressing the general N-polyhedron case. We propose a new vertex-centric view of the problem, which both simplifies the algorithm computing resulting geometric contributions, and vastly facilitates its spatial decomposition. We then embed the entire problem in a single KD-tree, specifically geared toward the final result by early pruning of any region of space not contributing to the final surface. This not only improves the robustness of the approach, it also gives it a fundamental speed advantage, with an output complexity depending on the output mesh size instead of the input size as with usual approaches. Complemented with a task-stealing parallelization, the algorithm achieves breakthrough performance, one to two orders of magnitude speedups with respect to state-of-the-art CPU algorithms, on boolean operations over two to several dozen polyhedra. The algorithm is also shown to outperform recent GPU implementations and approximate discretizations, while producing a topologically exact output without redundant facets. This algorithm was published as Inria research report [16].

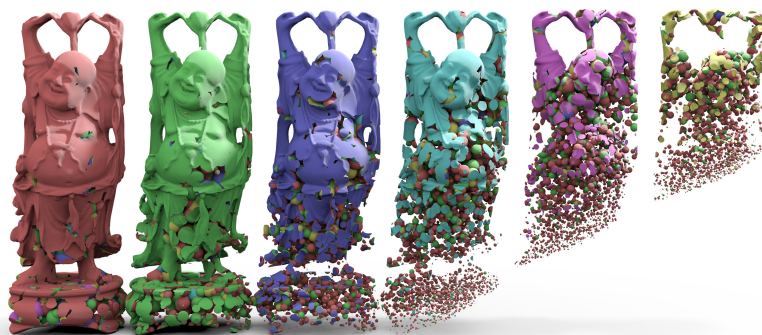


Figure 3. Intersection of 6 Buddhas with the union of 100,000 spheres (total 24 million triangles). Computed in 8 seconds on a desktop machine [16]

7.2. An Efficient Volumetric Framework for Shape Tracking

Recovering 3D shape motion using visual information is an important problem with many applications in computer vision and computer graphics, among other domains. Most existing approaches rely on surface-based strategies, where surface models are fit to visual surface observations. While numerically plausible, this paradigm ignores the fact that the observed surfaces often delimit volumetric shapes, for which deformations are constrained by the volume inside the shape. Consequently, surface-based strategies can fail when the observations define several feasible surfaces, whereas volumetric considerations are more restrictive with

respect to the admissible solutions. In this work, we investigate a novel volumetric shape parametrization to track shapes over temporal sequences. In contrast to Eulerian grid discretizations of the observation space, such as voxels, we consider general shape tessellations yielding more convenient cell decompositions, in particular the Centroidal Voronoi Tessellation. With this shape representation, we devise a tracking method that exploits volumetric information, both for the data term evaluating observation conformity, and for expressing deformation constraints that enforce prior assumptions on motion. Experiments on several datasets demonstrate similar or improved precisions over state-of-the-art methods, as well as improved robustness, a critical issue when tracking sequentially over time frames. This work was accepted as **oral** at CVPR 2015 (less than 3% acceptance rate) [8].

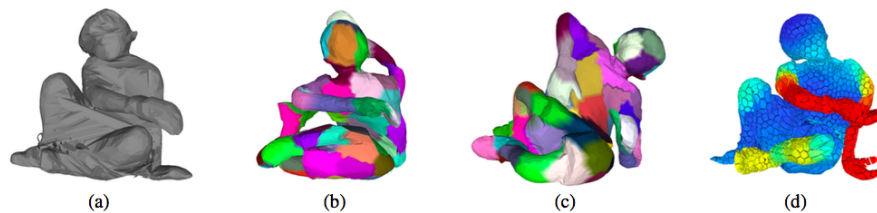


Figure 4. Frames of the GOALKEEPER dataset acquired on the Kinovis platform. (a) Visual hull input. (b) Tracking result of Cagniard et al. 2010. (c) Allain et al. 2014. (d) This method [8]. Note the improved angular shapes and the improved robustness.

7.3. Sparse Multi-View Consistency for Object Segmentation

Multiple view segmentation consists in segmenting objects simultaneously in several views. A key issue in that respect and compared to monocular settings is to ensure propagation of segmentation information between views while minimizing complexity and computational cost. In this work, we first investigate the idea that examining measurements at the projections of a sparse set of 3D points is sufficient to achieve this goal. The proposed algorithm softly assigns each of these 3D samples to the scene background if it projects on the background region in at least one view, or to the foreground if it projects on foreground region in all views. Second, we show how other modalities such as depth may be seamlessly integrated in the model and benefit the segmentation. The paper exposes a detailed set of experiments used to validate the algorithm, showing results comparable with the state of art, with reduced computational complexity. We also discuss the use of different modalities for specific situations, such as dealing with a low number of viewpoints or a scene with color ambiguities between foreground and background. This work was published as article in the PAMI journal [3].

7.4. Building Statistical Shape Spaces for 3D Human Modeling

Statistical models of 3D human shape and pose learned from scan databases have developed into valuable tools to solve a variety of vision and graphics problems. Unfortunately, most publicly available models are of limited expressiveness as they were learned on very small databases that hardly reflect the true variety in human body shapes. In this paper, we contribute by rebuilding a widely used statistical body representation from the largest commercially available scan database, and making the resulting model available to the community (visit <http://humanshape.mpi-inf.mpg.de>). As preprocessing several thousand scans for learning the model is a challenge in itself, we contribute by developing robust best practice solutions for scan alignment that quantitatively lead to the best learned models. We make implementations of these preprocessing steps also publicly available. We extensively evaluate the improved accuracy and generality of our new model, and show its improved performance for human body reconstruction from sparse input data. This work was published as Max Planck research report [17].



Figure 5. Three views of the PLANT dataset as processed by our method for multi-view silhouette extraction [3].

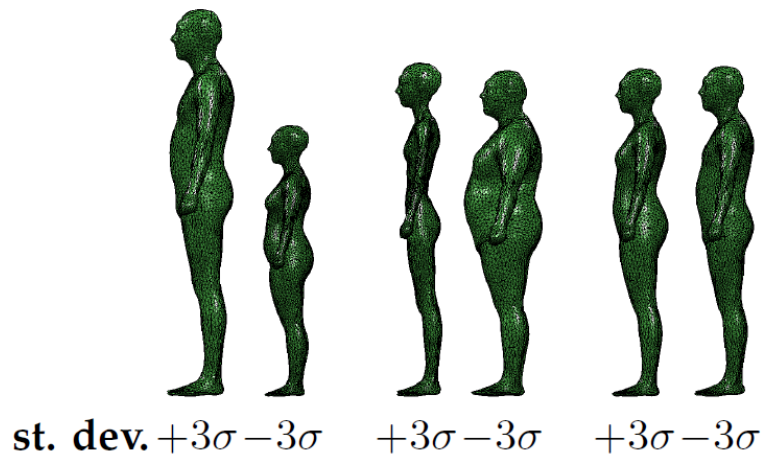


Figure 6. Visualization of the first three principal components learned from a large database of posture-normalized 3D human body scans [17].

7.5. A Groupwise Multilinear Correspondence Optimization for 3D Faces

Multilinear face models are widely used to model the space of human faces with expressions. For databases of 3D human faces of different identities performing multiple expressions, these statistical shape models decouple identity and expression variations. To compute a high-quality multilinear face model, the quality of the registration of the database of 3D face scans used for training is essential. Meanwhile, a multilinear face model can be used as an effective prior to register 3D face scans, which are typically noisy and incomplete. Inspired by the minimum description length approach, we propose the first method to jointly optimize a multilinear model and the registration of the 3D scans used for training. Given an initial registration, our approach fully automatically improves the registration by optimizing an objective function that measures the compactness of the multilinear model, resulting in a sparse model. We choose a continuous representation for each face shape that allows to use a quasi-Newton method in parameter space for optimization. We show that our approach is computationally significantly more efficient and leads to correspondences of higher quality than existing methods based on linear statistical models. This allows us to evaluate our approach on large standard 3D face databases and in the presence of noisy initializations. This work was published at the ICCV conference [9].

7.6. A statistical shape space model of the palate surface trained on 3D MRI scans of the vocal tract

We describe a minimally-supervised method for computing a statistical shape space model of the palate surface. The model is created from a corpus of volumetric magnetic resonance imaging (MRI) scans collected from 12 speakers. We extract a 3D mesh of the palate from each speaker, then train the model using principal component analysis (PCA). The palate model is then tested using 3D MRI from another corpus and evaluated using a high-resolution optical scan. We find that the error is low even when only a handful of measured coordinates are available. In both cases, our approach yields promising results. It can be applied to extract the palate shape from MRI data, and could be useful to other analysis modalities, such as electromagnetic articulography (EMA) and ultrasound tongue imaging (UTI). This work was published at the 18th International Congress of Phonetic Sciences [11].

7.7. Toward User-specific Tracking by Detection of Human Shapes in Multi-Cameras

Human shape tracking consists in fitting a template model to temporal sequences of visual observations. It usually comprises an association step, that finds correspondences between the model and the input data, and a deformation step, that fits the model to the observations given correspondences. Most current approaches find their common ground with the Iterative-Closest-Point (ICP) algorithm, which facilitates the association step with local distance considerations. It fails when large deformations occur, and errors in the association tend to propagate over time. In this paper, we propose a discriminative alternative for the association, that leverages random forests to infer correspondences in one shot. It allows for large deformations and prevents tracking errors from accumulating. The approach is successfully integrated to a surface tracking framework that recovers human shapes and poses jointly. When combined with ICP, this discriminative association proves to yield better accuracy in registration, more stability when tracking over time, and faster convergence. Evaluations on existing datasets demonstrate the benefits with respect to the state-of-the-art. This work was published at CVPR 2015 [12].

7.8. Video based Animation Synthesis with the Essential Graph

We propose a method to generate animations using video-based mesh sequences of elementary movements of a shape. New motions that satisfy high-level user-specified constraints are built by recombining and interpolating the frames in the observed mesh sequences. The interest of video based meshes is to provide real full shape information and to enable therefore realistic shape animations. A resulting issue lies, however, in the difficulty

to combine and interpolate human poses without a parametric pose model, as with skeleton based animations. To address this issue, our method brings two innovations that contribute at different levels: Locally between two motion sequences, we introduce a new approach to generate realistic transitions using dynamic time warping; More globally, over a set of motion sequences, we propose the essential graph as an efficient structure to encode the most realistic transitions between all pairs of input shape poses. Graph search in the essential graph allows then to generate realistic motions that are optimal with respect to various user-defined constraints. We present both quantitative and qualitative results on various 3D video datasets. They show that our approach compares favourably with previous strategies in this field that use the motion graph. This work was published at the 3DV 2015 conference [10].

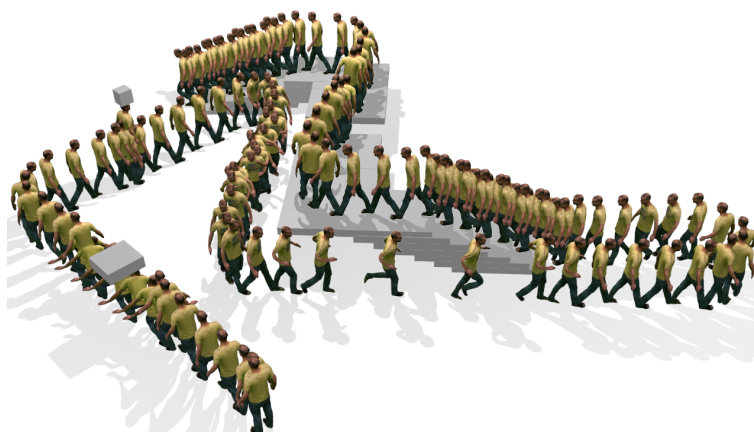


Figure 7. Example of 4D animation generated using by combining recorded 4D sequences [10].

7.9. Implicit B-Spline Surface Reconstruction

This paper presents a fast and flexible curve/surface reconstruction technique based on implicit b-spline. This representation does not require any parameterization and it is locally supported. This fact has been exploited in this paper to propose a reconstruction technique through solving a sparse system of equations. This method is further accelerated to reduce the dimension to the active control lattice. Moreover, the surface smoothness and user interaction are allowed for controlling the surface. Finally, a novel weighting technique has been introduced in order to blend small patches and smooth them in the overlapping regions. The whole framework is very fast and efficient and can handle large cloud of points with low computational cost. The experimental results show the flexibility and accuracy of the proposed algorithm to describe objects with complex topologies. Comparisons with other fitting methods highlight the superiority of the proposed approach in the presence of noise and missing data. This work was published as journal article in IEEE Transactions on Image Processing [6].

7.10. A Bayesian Approach to Multi-view 4D Modeling

This paper considers the problem of automatically recovering temporally consistent animated 3D models of arbitrary shapes in multi-camera setups. An approach is presented that takes as input a sequence of frame-wise reconstructed surfaces and iteratively deforms a reference surface such that it fits the input observations. This approach addresses several issues in this field that include: large frame-to-frame deformations, noise, missing data, outliers and shapes composed of multiple components with arbitrary geometries. The problem is cast as a geometric registration with two major features. First, surface deformations are modeled using

mesh decomposition into elements called patches. This strategy ensures robustness by enabling flexible regularization priors through inter-patch rigidity constraints. Second, registration is formulated as a Bayesian estimation that alternates between probabilistic data-model association and deformation parameter estimation. This accounts for uncertainties in the acquisition process and allows for noise, outliers and missing geometries in the observed meshes. In the case of marker-less 3D human motion capture, this framework can be specialized further with additional articulated motion constraints. Extensive experiments on various 4D datasets show that complex scenes with multiple objects of arbitrary nature can be processed in a robust way. They also demonstrate that the framework can capture human motion and provides visually convincing as well as quantitatively reliable human poses. This work was published as journal article in International Journal on Computer Vision (IJCV) [4].

7.11. A Hierarchical Approach for Regular Centroidal Voronoi Tessellations

In this paper we consider Centroidal Voronoi Tessellations (CVTs) and study their regularity. CVTs are geometric structures that enable regular tessellations of geometric objects and are widely used in shape modeling and analysis. While several efficient iterative schemes, with defined local convergence properties, have been proposed to compute CVTs, little attention has been paid to the evaluation of the resulting cell decompositions. In this paper, we propose a regularity criterion that allows us to evaluate and compare CVTs independently of their sizes and of their cell numbers. This criterion allows us to compare CVTs on a common basis. It builds on earlier theoretical work showing that second moments of cells converge to a lower bound when optimising CVTs. In addition to proposing a regularity criterion, this paper also considers computational strategies to determine regular CVTs. We introduce a hierarchical framework that propagates regularity over decomposition levels and hence provides CVTs with provably better regularities than existing methods. We illustrate these principles with a wide range of experiments on synthetic and real models.

This work was published as a journal article in Computer Graphics Forum [7].

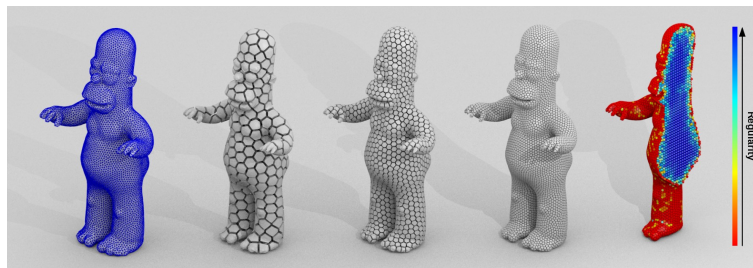


Figure 8. Hierarchical computation of a centroidal Voronoi tessellation from a 3D mesh [7]. Inside cells are very regular.

7.12. Just Noticeable Distortion Profile for Flat-Shaded 3D Mesh Surfaces

It is common that a 3D mesh undergoes some lossy operations (e.g., compression, watermarking and transmission through noisy channels), which can introduce geometric distortions as a change in vertex position. In most cases the end users of 3D meshes are human beings; therefore, it is important to evaluate the visibility of introduced vertex displacement. In this paper we present a model for computing a Just Noticeable Distortion (JND) profile for flat-shaded 3D meshes. The proposed model is based on an experimental study of the properties of the human visual system while observing a flat-shaded 3D mesh surface, in particular the contrast sensitivity function and contrast masking. We first define appropriate local perceptual properties on 3D meshes. We then detail the results of a series of psychophysical experiments where we have measured the threshold

needed for a human observer to detect the change in vertex position. These results allow us to compute the JND profile for flat-shaded 3D meshes. The proposed JND model has been evaluated via a subjective experiment, and applied to guide 3D mesh simplification as well as to determine the optimal vertex coordinates quantization level for a 3D model.

This work was published as a journal article in IEEE Transactions on Visualization and Computer Graphics [5].

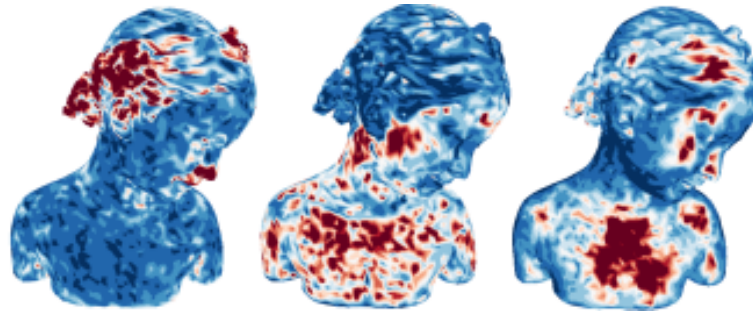


Figure 9. Just noticeable distortion profile in a light independent mode (left, middle) or with a light fixed in front of the model (right), for vertex displacements in the normal direction (left, right) or in the tangent direction (middle) [5].

MULTISPEECH Project-Team

7. New Results

7.1. Explicit Modeling of Speech Production and Perception

Participants: Yves Laprie, Slim Ouni, Vincent Colotte, Anne Bonneau, Agnès Piquard-Kipffer, Emmanuel Vincent, Denis Jouviet, Julie Busset, Benjamin Elie, Andrea Bandini, Ilef Ben Farhat, Sara Dahmani, Valérian Girard.

7.1.1. *Articulatory modeling*

7.1.1.1. *Acoustic simulations*

The acoustic simulation plays a key role in articulatory synthesis since it generates the acoustic signal from the instantaneous geometry of the vocal tract. This year we extended the single-matrix formulation to enable self-oscillation models of vocal folds, including glottal chinks, to be connected to the vocal tract. It also integrates the case of a local division of the main air path into two lateral channels, as it may occur during the production of lateral approximants. Extensions give rise to a reformulation of the acoustic conditions at the glottis, and at the upstream connection of bilateral channels. Numerical simulations validate the simulation framework. In particular the presence of a zero around 4 kHz due to the presence of bilateral channels around both sides of the tongue for the sound /l/ is confirmed by the simulations. These results agree with those obtained via independent techniques. Simulations of static vowels reveal that the behavior of the vocal folds is qualitatively similar whether they are connected to the single-matrix formulation or to the classic reflection type line analog model.

7.1.1.2. *Acquisition of articulatory data*

Magnetic resonance imaging (MRI) is a technique which provides very good static images of the vocal tract. However, it cannot be used directly to acquire dynamic images of the vocal tract which would enable a better comprehension of articulatory phenomena and the development of better coarticulation models. We thus have a cooperation with the IADI (Imagerie Adaptative Diagnostique et Interventionnelle) INSERM laboratory in Nancy Hospital intended to develop cineMRI [86], [87] (see. 6.8).

7.1.1.3. *Articulatory models*

An articulatory model of the velum [66], [65] was developed in order to complete an articulatory model already comprising other articulators. The velum contour was delineated and extracted from a thousand of X-ray images corresponding to short sentences in French. A principal component analysis was applied in order to derive the main deformation modes. The first component corresponds to the opening and comes with a shape modification linked to the apparition of a bulb in the upper part of the velum when it rises. The area function of the oral tract is modified so as to incorporate the velum movements. This model was connected with acoustic simulations in order to synthesize sentences containing French nasal vowels and consonants.

7.1.2. *Expressive acoustic-visual synthesis*

During this year, we have focused on the development of the acquisition infrastructure necessary to acquire audiovisual data. Mainly, we have developed several methods that allow acquiring acoustic and visual data synchronously. The visual data can originate from the Articulograph, Vicon or Intel RealSense devices. This heterogeneity of the data needs developing techniques to merge precisely the data in one unique reference. Synchronization techniques have also been developed for this purpose. We have evaluated the precision of the acquisition of such systems [61]. The combination of more than one motion capture technique aims to use the best quality data for each part of the face: (1) EMA (articulograph) for the lips, to have high precise measurement of the shape of the mouth that is related to speech and (2) kinect-like or Vicon system for the upper part of the face, that model mainly expressions.

We have acquired a small expressive audiovisual speech corpus of two actors: based on motion capture data (Vicon) and acoustic data. The content of the corpus is composed of six basic emotions (joy, sadness, anger, surprise, disgust and fear). This corpus will be used to investigate the characterization of emotions in audiovisual speech in the visual space and in the acoustic space.

We have also developed an algorithm to animate the 3D model of human face from a limited number of markers. The animation is very efficient and provides realistic animation results [82]. The 3D face will be used with the audiovisual system.

7.1.3. Categorization of sounds and prosody for native and non-native speech

7.1.3.1. Categorization of sounds for native speech

We investigated the schooling of a population of 166 students from primary to intermediate and secondary schools. These children and teenagers had specific language impairment: SLI (severe language impairment), dyslexia, dysorthographia. Since their childhood, they faced phonemic discrimination, phonological and phonemic analysis difficulties. We observed that they had trouble learning to read and more generally they experienced learning difficulties. Consequently, this lead them to repeat one or more grades, whereas in France, repetition is prohibited within each cycle and very limited between cycles.

7.1.3.2. Analysis of non-native pronunciations

Thanks to the detailed manual annotation of the French-German learner corpus that was carried out at the phonetic level in the IFCASL project (cf. 9.1.2), it was possible to investigate non-native pronunciation variants. The analysis revealed that German learners of French have most problems with obstruents in word-final position, whereas French learners of German show complex interferences with the vowel contrasts for length and quality [41]. Also, the correct pronunciation rate of the sounds, for several phonetic classes, was analyzed with respect to the learner's level, and compared to native pronunciations. One outcome is that different sound classes show different correct rates over the proficiency levels; and, for the German data, the frequently occurring syllabic [=n] is a prime indicator of the proficiency level.

We analyzed the realizations of French voiced fricatives by German non-native and French native speakers, in final position of an accentual group, a position where German fricatives are devoiced [27], [28]. Three speaker levels (from beginners to advanced) and different boundary types (depending on whether the fricative is followed by a pause, a schwa, or is directly followed by the first phoneme of the subsequent group) were considered. A set of cues, among which periodicity and fricative duration, have been analyzed. Results argue in favor of an influence of L1 (German) final devoicing on non-native realizations and show a strong interdependence between voicing, speakers' level, prosodic boundaries. The influence of orthography also strongly influenced voicing results.

We also investigated the realization of the short/long German contrast by French learners through three methods [60]. All these methods - phonetic annotation, perceptual experiment and acoustic analysis - used the same database (the IFCASL corpus). Depending on the method the results shed light on slightly different aspects of the same process, the interference of the French phonetic and phonological systems on the production of the German L2 vowels. Whereas the first method (phonetic annotation) revealed that especially rounded vowels are problematic in the long/short distinction, we could show with the second method (a perceptual experiment) that particularly the [o:]/[O] distinction seems to be hard to produce for French learners. The third method (an acoustical analysis) corroborated this finding and added acoustic details on duration and formants. The results of the studies can be used to create individualized training and feedback for foreign language learners, aimed at reducing their accent in L2.

7.2. Statistical Modeling of Speech

Participants: Antoine Liutkus, Emmanuel Vincent, Irina Illina, Dominique Fohr, Denis Jouviet, Joseph Di Martino, Vincent Colotte, Ken Deguernel, Amal Houdhek, Xabier Jaureguiberry, Aditya Nugraha, Luiza Orosanu, Imran Sheikh, Nathan Souviraà-Labastie, Dung Tran, Imene Zangar, Mohamed Bouallegue, Thibaut Fux, Emad Girgis, Juan Andres Morales Cordovilla, Sunit Sivasankaran, Freha Boumazza.

7.2.1. Source separation

Audio source separation is an inverse problem, which requires the user to guide the separation process using prior models for the source spectra and their spatial covariance matrices. We studied the impact of deterministic subspace constraints [14] over the spatial covariance matrices and pursued our work on the separation of multichannel mixtures guided by multiple, deformed reference signals such as repeated excerpts of the same music or repeated versions of the same sentence uttered by different speakers [17], [56]. Other models we have been working on include those based on local regularities of the spectral representations of musical sources (KAM, [52], [43], [51]). We also validated the positive impact of speech enhancement based on the FASST toolbox on speaker recognition [53].

As a new research direction, we extended the Gaussian framework for source separation to the family of α -stable stochastic processes [42]. This extension notably opens the path to new and robust parameters estimation algorithms for source separation [16], [67], that should be less prone to local minima. Current research notably comprises multichannel stable processes.

In parallel, we started yet another research track on the use of deep learning for source separation [24]. We proposed a new multichannel enhancement technique that exploits both the spatial properties of the sources as modeled by their spatial covariance matrices and their spectral properties as modeled by a deep neural network [75]. The model parameters are alternately estimated in an expectation-maximization (EM) fashion. We used this technique for music separation and speech enhancement in the context of the 2015 Signal Separation Evaluation Campaign (SiSEC) and the 3rd CHiME Speech Separation and Recognition Challenge, respectively [55]. We also used deep learning to address the fusion of multiple source separation techniques and found it to perform much better than the variational Bayesian model averaging techniques previously investigated [81].

Finally, we pursued our long-lasting efforts on the evaluation of audio source separation by co-organizing the 2015 Signal Separation Evaluation Campaign (SiSEC) [69] and writing a position paper about the scaling up of dataset sizes [21].

The ANR young researcher project KAMoulox (2016-2019 - cf. 9.1.5), that has just been accepted will deal with large audio archives, and more precisely with the "Archives du CNRS — Musée de l'homme" that gather a large set of old and noisy audio recordings (cf. 4.4). The work on source separation can lead to the design of semi automatic denoising and enhancement features, that would allow these researchers to significantly enhance their investigation capabilities, even without expert knowledge in sound engineering.

7.2.2. Acoustic modeling

We explored the use of an auxiliary function technique for fast training of neural networks [58]. We did not apply this technique to deep neural network acoustic models yet.

In the framework of using speech recognition for helping communication with deaf or hard-of-hearing people, robustness of the acoustic modeling was investigated. Studies were related to improving robustness with respect to speech signal level and environment noise through multicondition training and enhanced set of acoustic features (noise robust features or standard features after spectral noise subtraction) [37].

7.2.3. Linguistic modeling

7.2.3.1. Out-of-vocabulary proper name retrieval

Recognition of proper names (PN) is a challenging task in information retrieval in large audio/video databases. Proper names are semantically rich and are usually key to understanding the information contained in a document. Within the ContNomina project (cf. 9.1.3), we focus on increasing the vocabulary coverage of a speech transcription system by automatically retrieving proper names from contemporary text documents. We proposed methods that dynamically augment the automatic speech recognition system vocabulary, using lexical and temporal features in diachronic documents (documents that evolve over the time). Our work uses temporal context modeling to capture the lexical information surrounding proper names so as to retrieve out-of-vocabulary (OOV) proper names and increase the automatic speech recognition vocabulary.

We proposed new methods to retrieve OOV PNs relevant to an audio news document by using probabilistic topic models. We addressed retrieval of rare OOV PNs, which further improves the recall. Our proposed lexical context model improves the mean average precision of OOV PN retrieval [62]. We also proposed a two step approach for recognition of OOV PNs in an audio document. The first step retrieves OOV PNs relevant to an audio document using probabilistic topic models; and the second step uses a phonetic search for the target OOV PNs using a k -differences approximate string matching algorithm [63]. In [64], we discuss two specific phenomena, word frequency bias and loss of specificity, which affect the retrieval of OOV PNs using Latent Dirichlet Allocation (LDA) topic models. We studied different entity-topic models, which are extensions of LDA designed to learn relations between words, topics and PNs. We showed that our proposed methods of rare OOV PN and lexical context re-ranking improve the recall and the mean average precision for the LDA and the entity-topic models.

For OOV retrieval, we proposed the continuous space word representation using neural networks. This continuous vector representation (word embeddings) is learned from large amounts of unstructured text data. To model semantic and lexical context of proper names, different strategies of local context modeling were proposed [34], [33]. We studied OOV PN retrieval using temporal versus topic context modeling, different word representation spaces for word-level and document-level context modeling, and combinations of retrieval results [38]. We extended the previously proposed neural networks for word embedding models: the word vector representation proposed by Mikolov is enriched by an additional non-linear transformation. This model allows to better take into account lexical and semantic word relationships [39].

7.2.3.2. Adding words in a language model

A novel approach was proposed to add some new words in an existing n -gram language model, based on a similarity measure between the new words to be added and words already present in the language model [47]. Based on a small set of sentences containing the new words and on a set of n -gram counts containing the known words (known for the current language model), we search for known words which have the most similar neighbor distribution (of the few preceding and few following neighbor words) to the new words. The similar words are determined through the computation of KL divergences on the distribution of neighbor words. The n -gram parameter values associated to the similar words are then used to define the n -gram parameter values of the new words.

7.2.3.3. Selecting data for training a language model

Large vocabulary language models for speech transcription are usually trained from large amounts of textual data collected from various sources, which are more or less related to the target task. Selecting data that matches the target task was investigated in this context [46], this leads to a small reduction of the perplexity, and a smaller size of the resulting language model.

7.2.3.4. Music language modeling

Similarly to speech, music involves several levels of information, from the acoustic signal up to cognitive quantities such as composer style or key, through mid-level quantities such as a musical score or a sequence of chords. The dependencies between mid-level and lower- or higher-level information can be represented through acoustic models and language models, respectively. We pursued our pioneering work on music language modeling, with a particular focus on the modeling of long-term structure [12]. We also assessed the applicability of our prior work on joint modeling of note and chord sequences to new corpora of improvised jazz music, with the difficulty that these corpora are very small.

7.2.4. Speech generation by statistical methods

7.2.4.1. Pathological voice transformation

With respect to pathological voice processing, a competing approach to signal processing techniques consists in recognizing the pathological voice in order to transform it in a text version that can be re-synthesized. Such an approach is currently being experimented, and preliminary results are quite encouraging [15].

7.2.4.2. HMM-based synthesis

This year, we started working on HMM-based synthesis in the framework of a CMCU PHC project with ENIT (Engineer school at Tunis-Tunisia; cf. 9.3.2.2). Two topics will be explored by two PhD students. The first topic deals with the building of an Arabic corpora along with the analysis of linguistic features which are relevant for the HMM-based synthesis of the Arabic language. The second topic deals with improving the quality of the HMM-based synthesis system. In parallel, we started applying the HTS system (HMM-based Speech Synthesis System) to the French language.

7.3. Uncertainty Estimation and Exploitation in Speech Processing

Participants: Emmanuel Vincent, Odile Mella, Dominique Fohr, Denis Jouvét, Agnès Piquard-Kipffer, Baldwin Dumortier, Luiza Orosanu, Dung Tran, Sucheta Ghosh, Antoine Chemardin, Aghilas Sini.

7.3.1. Uncertainty and acoustic modeling

7.3.1.1. Noise-robust speech recognition

In many real-world conditions, the target speech signal overlaps with noise and some distortion remains after speech enhancement. In order to motivate further work by the community, we created an international evaluation campaign on that topic in 2011: the CHiME Speech Separation and Recognition Challenge. After two successful editions in 2011 and 2013, we organized the third edition in 2015 [25].

The framework of uncertainty decoding assumes that this distortion has a Gaussian distribution and seeks to estimate its covariance matrix in order to exploit it for subsequent feature extraction and decoding. A number of uncertainty estimators have been proposed in the literature, which are typically based on fixed mathematical approximations or heuristics. We made a conceptual breakthrough by proposing to learn the estimator from data using a non-parametric estimator and discriminative training [18], [59]. With GMM-HMM acoustic models, we obtained on the order of 30% relative word error rate reduction with respect to conventional decoding (without uncertainty), that is about twice as much as the reduction achieved by the best single uncertainty estimator. We also started working on the propagation of uncertainty in deep neural network acoustic models [19] and on its use for noise-robust speaker recognition [54].

7.3.1.2. Other applications

Besides the above applications, we started exploring applications of uncertainty modeling to robot audition [23] and control of wind turbines [31]. In the first context, uncertainty arises about the location of acoustic sources and the robot is controlled to locate the sources as quickly as possible. In the second context, uncertainty arises about the noise intensity of each wind turbine and the turbines are controlled to maximize electrical production under a maximum noise threshold.

7.3.2. Uncertainty and speech recognition

In the framework of using speech recognition for helping communication with deaf or hard-of-hearing people in the FUI project RAPSODIE (cf. 9.1.7), the best way for displaying the speech transcription results has been investigated. To our knowledge there is no suitable, validated and currently available display of the output of automatic speech recognizer for hard-of-hearing persons, in terms of size, colors and choice of the written symbols. The difficulty comes from the fact that speech transcription results contain recognition errors, which may impact the understanding process. Although the speech recognition system does not know the errors it makes, through the computation of confidence measures, the speech recognizer estimates if a word or a syllable is rather correctly recognized or not; hence such information can be used to adjust the display of the transcription results. Different ways were investigated for displaying the speech recognition results which take also into account the reliability of the recognized items. In this qualitative study, 10 persons have been interviewed to find the best way of displaying the speech transcription results. All the participants are deaf with different levels of hearing loss and various modes of communication [50].

7.3.3. Uncertainty and phonetic segmentation

Within the framework of the IFCASL project (cf. 9.1.2), a speech corpus of native and non-native speech for the French-German language pair was designed and recorded. Besides being used for analyzing non-native phenomena (cf. 7.1.3.2), this corpus will be used for developing and assessing automatic algorithms that will provide diagnosis on the learner mispronunciations [78]. Therefore, the automatic alignments of the audio files corresponding to the French and German speakers uttering French sentences (4100 audio files) were manually checked and corrected by a group of seven French annotators (the German data were handled by the German partner). We analyzed with CoALT the inter-annotator agreement with respect to an expert annotator for boundary shifts, insertions and deletions as well as devoicing diacritic [45]. The accuracy of the phone boundaries on non-native speech were investigated with respect to the HMM acoustic models used. The best performance (smallest amount of non-native phone segments whose boundaries are shifted by more than 20 ms compared to the manual boundaries) was obtained by combining each French native HMM model with an automatically selected German native HMM model [35].

Within the ANR ORFEO project (cf. 9.1.6), we addressed the problem of the alignment of spontaneous speech. The audio files processed in the ORFEO project were recorded under various conditions with a large SNR range and contain extra speech phenomena and overlapping speech. We trained several sets of acoustic models and tested different methods to adapt them to the various audio files [36]. Moreover in the framework of the EQUIPEX ORTOLANG (cf. 9.1.1), a web application, ASTALI (cf. 6.2), was developed in order to align a speech signal with its corresponding orthographic transcription (given in simple text file for short audio signals or in .trs files as generated by transcriber for longer speech signals).

In conventional speech-text alignments, a 10 ms frame shift is usually used for the acoustic analysis which leads to a minimum duration of 30 ms for each phone segment. Such duration constraint may not fit with actual sound duration in fast speaking rate. To overcome such constraint, a 5 ms frame shift can be used. Statistics on pronunciations variants estimated on large speech corpora have shown that when the conventional 10 ms frame shift is used, the frequency of the longest pronunciation variants gets underestimated [26]. Moreover, the analysis of some pronunciation variant frequencies have shown that some final consonantal cluster completely disappear at high speaking rates [40].

7.3.4. Uncertainty and prosody

Detection of sentence modality (question vs. affirmation) has been investigated using linguistic and prosodic features. Best results are achieved when the classifier uses all the available information [48], that is both linguistic and prosodic features. A detailed analysis has also shown that small errors in the determination of the sentence boundaries are not critical [49].

Speech-text alignments have been used to extract speech segments containing words and expressions that can be used either as normal lexical words or as discourse particles (as for example *quoi*, *voilà*, ...). The prosodic features for these words and expressions were extracted and analyzed [30]; automatic identification of the word function (discourse particle or not) from these prosodic features was also investigated.

In the context of the EQUIPEX ORTOLANG (cf. 9.1.1), several algorithms for computing the fundamental frequency have been implemented in the JSnoori software. These features can be computed directly from the GUI interface or through Python scripts. Future work will focus on improving the quality and robustness of the fundamental frequency estimation, and on determining the reliability of the estimations.

OAK Project-Team

7. New Results

7.1. Scalable and Expressive Techniques for the Semantic Web

On the topic of efficient query answering methods for semantic-rich RDF data, we have obtained new fundamental results for the RDF Schema ontology language [25] and for a simple DL-Lite dialect [23], [34]; we presented our results in a tutorial at IEEE ICDE [10] and in an invited keynote at SEBD, the Italian Database conference [4]. A demonstration issued from this work was presented at VLDB [26] and at BDA, the French database conference [27].

To help users get acquainted with large and complex RDF graphs, we have started to work on an approach for RDF graph summarization: a graph summary is a smaller RDF graph, often by several orders of magnitude, which preserves the core structural information of the original graph and thus allows to reason about several important graph property on a much more manageable structure. Our first results were presented in [17] and demonstrated at [29] and [30]. These results were also presented in the keynote of the Data Engineering and the Semantic Web workshop [5].

On the related topic of analytical RDF schemas, we have published novel techniques for incrementally computing the result of an RDF analytical query (also known as “RDF cube”) out of the result of a previously computed RDF cube [31]. Such computations, commonly known as roll-up, drill-down etc. in the classical relational database setting, require novel solutions for RDF due to the heterogeneity of the graph structure.

7.2. Massively Distributed Data Management Systems

One of the main results of the year is the publication of the full paper [15] and demonstration [14] on CliqueSquare in the highly prestigious IEEE Conference on Data Engineering (ICDE). CliqueSquare has also been released in open source in 2014 (see the Software section). Its main advantage is a novel technique for optimizing conjunctive queries in a massively parallel setting, using n-ary join operators; this allow the optimization algorithm to build plans which are as flat as possible. These results apply beyond the RDF conjunctive query evaluation to the general setting of relational conjunctive query processing in a massively parallel context.

Another crucial result of the year is the publication of the PAXQuery framework for massively processing XML queries based on the Stratosphere (now Apache Flink) platform [3]. We show that our algebra-based approach allows to capture the expressive processing performed by an XQuery query and to compile it efficiently into massively distributed plans which are then evaluated by the Flink platform; this outperforms a set of state-of-the-art approaches for evaluating XQuery queries in a parallel environment. The system was also demonstrated at SIGMOD [11].

7.3. Advanced Algorithms for Data Querying and Transformation

We focused on explaining why some data, so-called missing answers, are not part of the result of a query, even though a developer expects them to be there.

The query-based explanations we return during query analysis serve as the starting point for our query rewriting process. Indeed, knowing the condition combinations pruning data relevant to the missing answers significantly narrows the search space for eligible query rewritings as we can first focus on finding solutions that only affect these query conditions. To further prune the search space, our current solution applies a cost model for rewritings based on several criteria, including edit distance to the original query, or the number of side-effects (tuples additionally appearing in the result of the rewritten query that are not our original missing answers). To select the best solutions w.r.t. the different dimensions of our cost model, we compute and return the skyline over these. We have demonstrated a preliminary version of the proposed algorithm in [8]. This work is reported [7] and in the PhD thesis of K. Tzompanaki [1].

7.4. Social Data Management and Crowdsourcing

Some particular tasks such as annotating data or matching entities have traditionally been outsourced to human workers for many years. But the last few years have seen the rise of a new research field called crowdsourcing that aims at delegating a wide range of tasks to human workers. Crowd workers tend to make mistakes, so that redundant tasks are typically submitted to mitigate errors. As the crowd is a relatively expansive resource, we have worked on building formal frameworks to improve the efficiency of these processes.

Our research has been focused on two kinds of queries: boolean queries (asking the crowd to identify relevant items in a list, e.g., meals containing a specific ingredient), and ranking queries (asking the crowd to retrieve one or a few preferred items; e.g., ski resorts). We proposed new algorithms and heuristics improving the state of the art for boolean queries, and claimed the first algorithms for ranking queries (more specifically, for top-k and skyline queries) in the comparison framework [16].

We considered top-k query answering in social tagging systems, also known as folksonomies, a problem that requires a significant departure from existing, socially agnostic techniques. In a network-aware context, one can and should exploit the social links, which can indicate how users relate to the seeker and how much weight their tagging actions should have in the result build-up. Beyond explicit social links, we also focus on uncovering implicit, potentially richer relationships from user interactions and exploiting them to improve core functionality such as search. Specifically we considered as-you-type search in a social network, where results socially close to the user asking the query are more relevant, and proposed an efficient algorithm presenting, for any (increasingly longer) prefix of the query as the user types it, the k most relevant results [28].

ORPAILLEUR Project-Team

7. New Results

7.1. The Mining of Complex Data

Participants: Mehwish Alam, Aleksey Buzmakov, Victor Codocedo, Miguel Couceiro, Adrien Coulet, Esther Galbrun, Nicolas Jay, Florence Le Ber, Luis-Felipe Melo, Amedeo Napoli, Chedy Raïssi, Mohsen Sayed, My Thao Tang, Yannick Toussaint.

Keywords: formal concept analysis, relational concept analysis, pattern structures, pattern mining, association rule, graph mining, sequence mining, biclustering

Pattern mining and Formal Concept Analysis are suitable symbolic methods for KDDK, that may be used for real-sized applications. Global improvements are carried out on the scope of applicability, the ease of use, the efficiency of the methods, and on the ability to fit evolving situations. Accordingly, the team is extending these symbolic data mining methods for working on complex data (e.g. textual documents, biological, chemical or medical data), involving objects with multi-valued attributes (e.g. domains or intervals), n-ary relations, sequences, trees and graphs.

7.1.1. FCA and Variations: RCA, Pattern Structures and Biclustering

Advances in data and knowledge engineering have emphasized the needs for pattern mining tools working on complex data. In particular, FCA, which usually applies to binary data-tables, can be adapted to work on more complex data. In this way, we have contributed to two main extensions of FCA, namely Pattern Structures and Relational Concept Analysis. Pattern Structures (PS [92]) allow to build a concept lattice from complex data, e.g. numbers, sequences, trees and graphs. Relational Concept Analysis (RCA) is able to analyze objects described both by binary and relational attributes [101] and can play an important role in text classification and text mining. Following this way, and regarding itemset and association rule discovery, we improved standard algorithms for building lattices from large data and for completing the algorithm collection of the Coron platform [103].

Many developments were carried out in pattern mining and FCA for improving data mining algorithms and their applicability, and for solving some specific problems such as information retrieval, discovery of functional dependencies and biclustering. We designed new information retrieval methods based on FCA where the concept lattice is considered as an index space for answering disjunctive queries [54]. We developed also a whole line of work on pattern structures for the discovery of functional dependencies [80], text classification and heterogeneous pattern structures [83], and pattern structures for structured attribute sets [46]. FCA can also be considered as a clustering method and we adapted pattern structures to clustering for analyzing numerical datatables supporting recommendation problems [13]. Projections can be associated with pattern structures for leveraging the volume and the complexity of the computation [53]. We designed also a quasi-polynomial algorithm for mining top patterns w.r.t. measures satisfying special properties in a FCA framework [52]. We also proposed new visualization techniques and tools able to display important and useful information (e.g. stable concepts) from large concept lattices [49].

Still considering complex data, we worked on the analysis of molecular structures (or molecular graphs) [34]. The mining of molecular graphs is an important task for many reasons, among which the challenges it represents regarding knowledge discovery, life sciences and healthcare, and, as well, the industrial needs that can be met whenever substantial results are obtained (especially in pharmacology).

7.1.2. Text Mining

Ontologies help software and human agents to communicate by providing shared and common domain knowledge, and by supporting various tasks, e.g. problem-solving and information retrieval. In practice, building an ontology or at least “ontological concept definitions” depends on a number of ontological resources having different types: thesaurus, dictionaries, texts, databases, and ontologies themselves. We are currently working on the design of a methodology based on FCA and RCA for ontology engineering from heterogeneous ontological resources. This methodology is based on both FCA and RCA, and was previously successfully applied in domains such as astronomy and biology.

In the framework of the ANR Hybride project (see 8.2.1.2), an engineer is implementing a robust system based on these previous research results, for preparing the way to new research directions involving trees and graphs. Moreover, we led a first successful experiment on extracting drug-drug interactions applying “lazy pattern structure classification” to syntactic trees [66]. In addition, in his thesis work, Mohsen Sayed focused on extracting relations between named entities using graph mining methods applied to dependency graphs. We are currently investigating how this approach can be generalized, i.e. how to detect a relation between complex expressions which are not previously recognized as named entities [64].

The notion of “Jumping Emerging Patterns” (JEP) previously used in chemistry [12], was updated and adapted to the context of text mining within the ANR Termith project. The objective is to design a learning method for filtering candidate terms within a full text and to decide whether an occurrence should be tagged as a term, i.e. as a positive example, or as a simple word, i.e. as a negative example. The method extracts from a training set all JEPs which are considered as hypotheses [7]. To reduce the number of JEPs and to only retain the most significant from a linguistic point of view, JEPs are weighted and a constraint solver is used to check the maximal coverage of the positive examples. Results are currently under evaluation.

7.1.3. Mining Sequences and Trajectories

Sequence data is widely used in many applications. Computing the similarity between sequences is a very important challenge for many different data mining tasks. There is a plethora of similarity measures for sequences in the literature, most of them being designed for sequences of items. In a recent work with Elias Egho, we study the problem of measuring the similarity between sequences of itemsets [32]. We focus on the notion of common subsequences as a way to measure similarity between a pair of sequences composed of a list of itemsets. In this work, we present new combinatorial results for efficiently counting distinct and common subsequences. These theoretical results are the cornerstone of an effective dynamic programming approach to deal with this problem. In addition, we develop an approximate method to speed up the computation process for long sequences. We have applied the method to various data sets: healthcare trajectories, on-line handwritten characters and synthetic data. The results confirm that the current similarity measure produces competitive scores and indicate that the method is relevant for large scale sequential data analysis.

Nowadays data sets are available in very complex and heterogeneous ways. Mining of such data collections is essential to support many real-world applications ranging from healthcare to marketing. In a recent work, we focused on the analysis of “complex sequential data” by means of interesting sequential patterns [19]. We approach the problem using FCA and pattern structures, where the subsumption relation ordering patterns is defined w.r.t. the partial order on sequences. We show how pattern structures along with projections, i.e. a data reduction of sequential structures, are able to enumerate more meaningful patterns and increase the computing efficiency of the approach. Finally, we demonstrate the applicability of the method for discovering and analyzing patient patterns from a French healthcare data set on cancer. The quantitative and qualitative results –with annotations and analysis from a physician– are reported in this use case which is one main motivation for this work.

7.1.4. Mining with Preferences

In the last decade, the pattern mining community has witnessed a sharp shift from efficiency-based approaches to methods which can extract more meaningful patterns. Recently, new methods adapting results from studies of economic efficiency and multi-criteria decision analysis such as Pareto efficiency, or skylines, have been studied. Within pattern mining, this novel line of research allows the easy expression of preferences according

to a dominance relation. We have developed approaches that are useful from a user-preference point of view, tending to promote the use of pattern mining algorithms for non-experts. These approaches are based on the discovery of skyline patterns, or “skypatterns”, in relation with condensed representations of patterns. This last relationship facilitates the computation of skypatterns, providing a flexible and efficient approach to mine skypatterns reusing a dynamic constraint satisfaction problems (CSP) framework [8].

7.1.5. Aggregation

Aggregation or consensus theory studies any process dealing the merging of several objects (numerical values, qualitative data, preferences, etc.) into a single (or several) object of similar type and that, in some way, is the best representation. The need to aggregate objects in a meaningful way has become more and more present in an increasing number of areas not only of mathematics, statistics or physics, but especially in applied fields such as engineering, computer science, social sciences and biology. In social choice and multicriteria decision aid, objects are preferences that are expressed by users, voters or criteria, and are modeled by order relations or utility functions. In cluster analysis, the objects to merge are classifications (such as partitions, hierarchies or trees) or related functions (such as similarity/dissimilarity measures).

With the proliferation of massive databases and new fields such as computational advertising, search engines and recommender systems, the need for information retrieval and knowledge discovery processes became emergent as well as the construction of user preference models for classification and prediction purposes. Also in biology and phylogenetics, aggregation is used to find consensus patterns among DNA sequences or finding consensus trees within taxonomies. As algorithms are often heuristic in such large datasets, they rarely produce the same output, highlighting the importance of finding means of aggregation to produce consensus structures. The difficulty in extracting such consensus structures comes down to define appropriate aggregation rules (e.g., counting and median procedures), and their impossibility is many times revealed by Arrowian results. A way to avoid such impossibility results is the consideration of alternative aggregation rules or the weakening of underlying structures, for instance weak hierarchies that allow overlapping clusters while keeping desirable tree-like properties.

We are working on a theoretical basis of a unified theory of consensus and to set up a general machinery for the choice and use of aggregation functions. This choice depends on properties specified by users or decision makers, the nature of the objects to aggregate as well as computational limitations due to prohibitive algorithmic complexity. This problem demands an exhaustive study of aggregation functions that requires an axiomatic treatment and classification of aggregation procedures as well as a deep understanding of their structural behavior. Moreover, Arrowian results are also envisioned since they constitute an important tool in the identification of reasonable algebraic/relational structures for representing data as well as in the identification of meaningful aggregation processes.

Direct applications of this theory are preference learning and cluster analysis. In the first case, preferences are represented by global utility functions and alternatives with higher utilities are preferred. Moreover, simplified versions of this model will be explored in the context of feature selection for both dimension reduction of data as well as classifier design. In the second case, we consider median structures that include several ordered/relational structures (trees, graphs, orders) and that allow several consensus procedures. This is particularly useful in a context of classification that takes into account evolutionary relations between classes, for instance, in taxonomical biology and phylogenetics.

7.1.6. Video Game Analytics

The video game industry has enormously grown over the last twenty years, bringing new challenges to the artificial intelligence and data analysis communities. We are studying the automatic discovery of strategies in real-time strategy games through pattern mining. Such patterns are the basic units for many tasks such as automated agent design, but also to build tools for the professionally played video games in the electronic sports scene. Continuing our joint collaboration with researchers from the MIT GameLab we successfully extended our previous work to a journal paper that will be published in 2016.

7.2. Knowledge Discovery in Healthcare and Life Sciences

Participants: Miguel Couceiro, Adrien Coulet, Amedeo Napoli, Chedy Raïssi, Mohsen Sayed, Malika Smaïl-Tabbone, Yannick Toussaint.

Life Sciences constitute a challenging domain for KDDK. Biological data are complex from many points of views, e.g. voluminous, high-dimensional and deeply inter-connected. Analyzing such data is a crucial issue in healthcare, environment and agronomy. Besides, many bio-ontologies are available and can be used to enhance the knowledge discovery process. Accordingly, the research work of the Orpailleur team in KDDK applied to Life Sciences is in concern with the use of bio-ontologies to improve KDDK, and as well information retrieval, access to “Linked Open Data” (LOD) and data integration.

7.2.1. Ontology-based Clustering of Biological Linked Open Data

Increasing amounts of biomedical data provided as Linked Open Data (LOD) offer novel opportunities for knowledge discovery in bio-medicine. We proposed an approach for selecting, integrating, and mining LOD with the goal of discovering genes responsible for a disease [99]. We are currently working on the integration of LOD about known phenotypes and genes responsible for diseases along with relevant bio-ontologies. We are also defining a corpus-based semantic distance. One possible application of this work is to build and compare possible diseaseomes, i.e. global graphs representing all diseases connected according to their pairwise similarity values.

7.2.2. Suggesting Valid Pharmacogenes by Mining Linked Open Data and Electronic Health Records

A standard task in pharmacogenomics research is identifying genes that may be involved in drug response variability and called “pharmacogenes”. As genomic experiments in this domain tend to generate many false positives, computational approaches based on background knowledge may generate more valuable results. Until now, the later have used only molecular networks databases or biomedical literature. We are studying and working on a novel method that take advantage of an eclectic set of linked data sources to validate uncertain drug–gene relationships, i.e. pharmacogenes [3]. One advantage relies on the standard implementation of linked data that facilitates the joint use of various sources and makes easier the consideration of features of various origins. Accordingly, we proposed an initial selection of linked data sources relevant to pharmacogenomics. We formatted these data to train a random forest algorithm, producing a model that classify drug–gene pairs as related or not, thus validating candidate pharmacogenes.

With this same motivation of validating state-of-the-art knowledge in pharmacogenomics, a new ANR project called “PractiKPharma” will be initiated in 2016 and will rely on similar ideas. The originality of “PractiKPharma” is to use “Electronic Health Records” to constitute cohorts of patients that are then mined for validating extracted pharmacogenomics knowledge units (<http://praktikpharma.loria.fr/>).

7.2.3. Biological Data Aggregation for Knowledge Discovery

During this year, in collaboration with the Capsid Team, we contributed to write up two multi-disciplinary projects with a group of clinicians from the Regional University Hospital (CHU Nancy) and bio-statisticians from the Maths Lab (IECL). The first project, entitled ITM2P⁰ lying in the so-called CPER 2015–2020 framework, was accepted and granted. The funding is mainly intended for medical and computing equipments and will be used to set up four scientific platforms. We are involved in the SMEC platform as a support for “Simulation, Modeling and Knowledge Extraction from Bio-Medical Data”.

The second project is a RHU⁰ project entitled *Fight Heart Failure* (FHF) and was accepted as a so-called “investissement d’avenir” and granted. We are in charge of a workpackage which will give us the opportunity of exploring important research questions. Among these questions, one is to define “data aggregation” mechanisms with a twofold objective: (i) the definition of pairwise patient similarity given that patients are described by complex dimensions involving relations and time and (ii) the efficient clustering of patients based

⁰“Innovations Technologiques, Modélisation et Médecine Personnalisée”

⁰“Recherche Hospitalo-Universitaire”

on this similarity measure. Each cluster should correspond to a bioprofile, i.e. a subgroup of patients sharing the same form of the disease and thus the same diagnosis and care strategy. For doing that, we are currently investigating consensus theories [95] and their applicability to a bio-medical context, and as well aggregation operators as defined in various contexts, e.g. databases, data-warehouses, web of data, and graph theory. The idea is to consider relational and temporal data aggregation as a first class citizen in the data preparation phase of the knowledge discovery. This allows to assess the contribution of aggregation for such a task and in this context.

Another question is related to the construction of a prediction model for each bioprofile/subgroup –once validated by the clinicians– to be used in a decision support system. This will likely require the combination of symbolic and numerical methods for the classification task.

7.2.4. Analysis of biomedical data annotated with ontologies

Annotating data with concepts of an ontology is a common practice in the biomedical domain. Resulting annotations define links between data and ontologies that are key for data exchange, data integration and data analysis. Since 2011, we collaborate with the National Center for Biomedical Ontologies (NCBO) to develop a large repository of annotations named the NCBO Resource Index. This repository contains annotations of 36 biomedical databases annotated with concepts of more than 200 ontologies of the BioPortal (<http://biportal.bioontology.org/>). In the preceding years, we compared the annotations of a database of biomedical publications (Medline) with two databases of scientific funding (Crisp and ResearchCrossroads) to profile disease research. One main challenge is to mine these annotations.

As a first attempt, we adapted pattern structures to analyze the annotations of biomedical databases [85]. We considered annotated biomedical documents as objects and the corresponding annotations were classified according to various dimensions, i.e. a particular aspect of domain knowledge. The resulting classification of annotations allowed not only to discover correlations between annotations but also incomplete annotations that could be fixed afterward. This adaptation of pattern structures opens many perspectives in term of ontology reengineering and knowledge discovery.

7.3. Knowledge Engineering and Web of Data

Participants: Mehwish Alam, Aleksey Buzmakov, Victor Codocedo, Emmanuelle Gaillard, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer.

Keywords: knowledge engineering, web of data, classification-based reasoning, case-based reasoning, belief revision, semantic web

7.3.1. Around the Taaable Research Project

The Taaable project was originally created as a challenger of the Computer Cooking Contest (ICCBR Conference) [84] (<http://intoweb.loria.fr/taaaable3ccc/>). Beyond its participation to the CCC challenges, the Taaable project aims at federating various research themes: case-based reasoning (CBR), information retrieval, knowledge acquisition and extraction, knowledge representation, minimal change theory, ontology engineering, semantic wikis, text-mining, etc. CBR performs adaptation of recipes w.r.t. user constraints. The reasoning process is based on a cooking domain ontology (especially hierarchies of classes) and adaptation rules. The knowledge base is encoded within a semantic wiki containing the recipes, the domain ontology and adaptation rules.

As acquiring knowledge from experts is costly, a new approach was proposed to allow a CBR system to use partially reliable, non expert, knowledge from the Web for reasoning. This approach is based on notions such as belief, trust, reputation and quality, as well as their relationships and rules to manage the knowledge reliability. The reliability estimation is used to filter knowledge with high reliability as well as to rank the results produced by the CBR system. Performing CBR with knowledge resulting from an e-community is improved by taking into account the knowledge reliability [61].

Another study shows how the case retrieval of a CBR system can be improved using typicality. Typicality discriminates subclasses of a class in the domain ontology depending of how a subclass is a good example for its class. An approach has been proposed to partition the subclasses of some classes into atypical, normal and typical subclasses in order to refine the domain ontology. The refined ontology allows a finer-grained generalization of the query during the retrieval process, improving at the same time the final results of the CBR system [62].

The Taaable system also includes a module for adapting textual preparations (from a source recipe text to an adapted recipe text, through a formal representation in the qualitative algebra INDU). The evaluation of this module as a whole thanks to users has been carried out and has shown its efficiency (w.r.t. text quality and recipe quality), when compared with another approach to textual adaptation [4].

FCA allows to organize objects according to the properties they share into a concept lattice. A lattice has been built on a large set a cooking recipes according to the ingredients they use, producing a hierarchy of ingredient combinations. When a recipe R has to be adapted, this lattice can be used to search the best ingredient combinations in the concepts that are the closest to the concept representing R [63].

Minimal change theory and belief revision can be used as tools to support adaptation in CBR, i.e. the source case is modified to be consistent with the target problem using a revision operator. Belief revision was applied to Taaable to adjust the ingredient quantities using engines included in the Revisor library (see § 6.4.5). This year, a mixed linear optimization has implemented to produce human easy understandable quantities. For example, when the ingredient is a lemon, its quantity will take the form of a quarter, a half, etc., instead of 54 g (which corresponds to a half lemon) [63].

7.3.2. Exploring and Classifying the Web of Data

A part of the research work in Knowledge Engineering is oriented towards knowledge discovery in the web of data, as, with the increased interest in machine processable data, more and more data is now published in RDF (Resource Description Framework) format. The popularization and quick growth of Linked Open Data (LOD) has led to challenging aspects regarding quality assessment and data exploration of the RDF triples that shape the LOD cloud. Particularly, we are interested in the completeness of the data and the their potential to provide concept definitions in terms of necessary and sufficient conditions [1]. We have proposed a novel technique based on Formal Concept Analysis which organizes subsets of RDF data into a concept lattice. This allows data exploration as well as the discovery of implication rules which are used to automatically detect missing information and then to complete RDF data and to provide definitions. Moreover, this is also a way of reconciling syntax and semantics in the LOD cloud. Experiments on the DBpedia knowledge base shows that this kind of approach is well-founded and effective.

Other important aspects are concerned with data access, data visualization w.r.t. the SPARQL query language [46], [49]. SPARQL queries over the web of data usually produce lists of tuples as answers that may be voluminous and hard to interpret. We introduced Lattice-Based View Access (LBVA), a framework based on FCA, which provides a classification of the answers of SPARQL queries based on a concept lattice. This concept lattice can be considered as a materialized view of the data resulting from a SPARQL query and can be navigated for retrieving or mining specific patterns. We associate a VIEW-BY clause to SPARQL for facilitating the interaction between analysts and LOD. The organization of answers is based on an original proposition on pattern structures for structured sets of attributes, which appears to be quite efficient and very well-adapted to the classification and analysis of RDF data. The visualization and the navigation of the concept lattice are guided by RV-Xplorer (i.e. RDF View eXplorer), an adapted interactive visualization system. Experiments show that the approach is well-founded and that it opens many new perspectives in the domain.

7.4. Advances in Graph Theory

Participants: Miguel Couceiro, Amedeo Napoli, Chedy Raïssi, Jean-Sébastien Sereni, Mario Valencia.

Keywords: graph theory, extremal graph theory, chromatic number, triangle-free graph, planar graph, graph coloring

We announced in the last report that we started to work on a conjecture by Heckman and Thomas from 1999. We managed to confirm the conjecture and the demonstration was published in January 2014. A classical result by Staton, from 1979, states that every triangle-free graph G with maximum degree at most 3 contains an independent set of order at least $5n/14$, where n is the number of vertices of G . Heckman and Thomas conjectured a stronger fact: the fractional chromatic number of such a graph is at most $14/5$. We confirmed their conjecture by establishing the following stronger assertion: for any assignment of weights (i.e., real numbers) to the vertices of such a graph G , there exists an independent set I such that the weights of the vertices in I is at least $5/14$ times the total weight of the G .

Exploring further the methods we introduced to solve this conjecture, we obtained new results concerning the fractional chromatic number of planar triangle-free graphs. While the fractional chromatic number of such graphs is at most 3 (because their chromatic number is), a construction of Jones proved the existence of triangle-free planar graphs with fractional chromatic number arbitrarily close to 3. Thus one wonders whether there could be such graphs with fractional chromatic number exactly 3. We demonstrated this not to be the case, by proving a general upper bound of $\frac{9n}{3n+1} = 3(1 - \frac{1}{3n+1})$ for every triangle-free planar graph G with n vertices. This bound is qualitatively the best possible: Jones's construction yields graphs with fractional chromatic number $3 - \frac{c}{n}$ for some constant c . In addition, a tight bound was obtained if the graphs considered are furthermore required to have maximum degree at most 4. In this case, the bound becomes $\frac{3n}{3n+1}$.

Motivated by frequency assignment in office blocks, we study the chromatic number of the adjacency graph of a 3-dimensional parallelepiped arrangement. In the case each parallelepiped is within one floor, a direct application of the Four-Colour Theorem yields that the adjacency graph has chromatic number at most 8. We provide an example of such an arrangement needing exactly 8 colors. We also discuss bounds on the chromatic number of the adjacency graph of general arrangements of 3-dimensional parallelepipeds according to geometrical measures of the parallelepipeds (side length, total surface area or volume).

PANAMA Project-Team

7. New Results

7.1. Recent results on sparse representations

Sparse approximation, high dimension, scalable algorithms, dictionary design, sample complexity

The team has had a substantial activity ranging from theoretical results to algorithmic design and software contributions in the field of sparse representations, which is at the core of the ERC project PLEASE (projections, Learning and Sparsity for Efficient Data Processing, see Section 9.2.1.1).

7.1.1. Theoretical results on sparse representations, graph signal processing, and dimension reduction

Participants: Rémi Gribonval, Yann Traonmilin, Gilles Puy, Nicolas Tremblay, Pierre Vandergheynst.

Main collaboration: Mike Davies (University of Edinburgh), Pierre Borgnat (ENS Lyon),

Stable recovery of low-dimensional cones in Hilbert spaces: Many inverse problems in signal processing deal with the robust estimation of unknown data from underdetermined linear observations. Low dimensional models, when combined with appropriate regularizers, have been shown to be efficient at performing this task. Sparse models with the ℓ_1 -norm or low rank models with the nuclear norm are examples of such successful combinations. Stable recovery guarantees in these settings have been established using a common tool adapted to each case: the notion of restricted isometry property (RIP). This year, we established generic RIP-based guarantees for the stable recovery of cones (positively homogeneous model sets) with arbitrary regularizers. These guarantees were illustrated on selected examples. For block structured sparsity in the infinite dimensional setting, we used the guarantees for a family of regularizers which efficiency in terms of RIP constant can be controlled, leading to stronger and sharper guarantees than the state of the art. A journal paper is currently under revision [57].

Recipes for stable linear embeddings from Hilbert spaces to \mathbb{R}^m : We considered the problem of constructing a linear map from a Hilbert space (possibly infinite dimensional) to \mathbb{R}^m that satisfies a restricted isometry property (RIP) on an arbitrary signal model set. We obtained a generic framework that handles a large class of low-dimensional subsets but also *unstructured* and *structured* linear maps. We provided a simple recipe to prove that a random linear map satisfies a general RIP on the model set with high probability. We also described a generic technique to construct linear maps that satisfy the RIP. Finally, we detailed how to use our results in several examples, which allow us to recover and extend many known compressive sampling results. This has been presented at the conference EUSIPCO 2015 [28], and a journal paper has been submitted [55].

Random sampling of bandlimited signals on graphs: We studied the problem of sampling k -bandlimited signals on graphs. We proposed two sampling strategies that consist in selecting a small subset of nodes at random. The first strategy is non-adaptive, i.e., independent of the graph structure, and its performance depends on a parameter called the graph coherence. On the contrary, the second strategy is adaptive but yields optimal results. Indeed, no more than $O(k \log(k))$ measurements are sufficient to ensure an accurate and stable recovery of all k -bandlimited signals. This second strategy is based on a careful choice of the sampling distribution, which can be estimated quickly. Then, we proposed a computationally efficient decoder to reconstruct k -bandlimited signals from their samples. We proved that it yields accurate reconstructions and that it is also stable to noise. Finally, we conducted several experiments to test these techniques. A journal paper has been submitted [56].

Accelerated spectral clustering: We leveraged the proposed random sampling technique to propose a faster spectral clustering algorithm. Indeed, classical spectral clustering is based on the computation of the first k eigenvectors of the similarity matrix' Laplacian, whose computation cost, even for sparse matrices, becomes prohibitive for large datasets. We showed that we can estimate the spectral clustering distance matrix without computing these eigenvectors: by graph filtering random signals. Also, we took advantage of the stochasticity of these random vectors to estimate the number of clusters k . We compared our method to classical spectral clustering on synthetic data, and show that it reaches equal performance while being faster by a factor at least two for large datasets. A conference paper has been accepted at ICASSP 2016 [43] and a long version is in preparation.

7.1.2. Algorithmic and theoretical results on dictionary learning

Participants: Rémi Gribonval, Luc Le Magoarou, Nicolas Bellot, Thomas Gautrais, Nancy Bertin, Srdan Kitic.

Main collaboration (theory for dictionary learning): Rodolphe Jenatton, Francis Bach (Equipe-projet SIERRA (Inria, Paris)), Martin Kleinstueber, Matthias Seibert (TU-Munich),

Theoretical guarantees for dictionary learning : An important practical problem in sparse modeling is to choose the adequate dictionary to model a class of signals or images of interest. While diverse heuristic techniques have been proposed in the litterature to learn a dictionary from a collection of training samples, there are little existing results which provide an adequate mathematical understanding of the behaviour of these techniques and their ability to recover an ideal dictionary from which the training samples may have been generated.

Beyond our pioneering work [86], [109] [5] on this topic, which concentrated on the noiseless case for non-overcomplete dictionaries, we showed the relevance of an ℓ^1 penalized cost function for the locally stable identification of overcomplete incoherent dictionaries, in the presence of noise and outliers [19]. Moreover, we established sample complexity bounds of dictionary learning and other related matrix factorization schemes (including PCA, NMF, structured sparsity ...) [20].

Learning computationally efficient dictionaries Classical dictionary learning is limited to small-scale problems. Inspired by usual fast transforms, we proposed a general dictionary structure that allows cheaper manipulation, and an algorithm to learn such dictionaries –and their fast implementation. The principle and its application to image denoising appeared at ICASSP 2015 [33] and an application to speedup linear inverse problems was published at EUSIPCO 2015 [32]. A journal paper is currently under revision [51].

We further explored the application of this technique to obtain fast approximations of Graph Fourier Transforms – a conference paper on this latter topic has been accepted for publication in ICASSP 2016 [41]. A C++ software library is in preparation to release the resulting algorithms.

Operator learning for cospase representations: Besides standard dictionary learning, we also considered learning in the context of the cospase model. The overall problem is to learn a low-dimensional signal model from a collection of training samples. The mainstream approach is to learn an overcomplete dictionary to provide good approximations of the training samples using sparse synthesis coefficients. This famous sparse model has a less well known counterpart, in analysis form, called the cospase analysis model. In this new model, signals are characterized by their parsimony in a transformed domain using an overcomplete analysis operator.

This year we obtained an upper bound of the sample complexity of the learning process for analysis operators, and designed a stochastic gradient descent (SGD) method to efficiently learn analysis operators with separable structures. Numerical experiments were provided that link the sample complexity to the convergence speed of the SGD algorithm. A journal paper has been published [24].

7.1.3. An alternative framework for sparse representations: analysis sparse models

Participants: Rémi Gribonval, Nancy Bertin, Srdan Kitic, Laurent Albera.

In the past decade there has been a great interest in a synthesis-based model for signals, based on sparse and redundant representations. Such a model assumes that the signal of interest can be composed as a linear combination of *few* columns from a given matrix (the dictionary). An alternative *analysis-based* model can be envisioned, where an analysis operator multiplies the signal, leading to a *cosparse* outcome. Building on our pioneering work on the cosparse model [101], [85], [102] successful applications of this approach to sound source localization, audio declipping and brain imaging have been developed this year.

Versatile co-sparse regularization: Digging the groove of last year results (comparison of the performance of several cosparse recovery algorithms in the context of sound source localization [94], demonstration of its efficiency in situations where usual methods fail ([96], see paragraph 7.5.2), applicability to the hard declipping problem [95], application to EEG brain imaging [60] (see paragraph 7.5.3), a journal paper embedding the latest algorithms and results in sound source localization and brain source localization in a unified fashion was published in IEEE Transactions on Signal Processing [23]. Other communications were made in conferences and workshops [50], [31] and Srđan Kitić defended his PhD thesis [12]. New results include experimental confirmation of robustness and versatility of the proposed scheme, and of its computational merits (convergence speed increasing with the amount of data)

Parametric operator learning for cosparse calibration: In many inverse problems, a key challenge is to cope with unknown physical parameters of the problem such as the speed of sound or the boundary impedance. In the sound source localization problem, we showed that the unknown speed of sound can be learned jointly in the process of cosparse recovery, under mild conditions (work presented last year at iTwist'14 workshop [66]). This year, improved and extended results were obtained: first with a new algorithm for sound source localization with unknown speed of sound [12], then by extending the formulation to the case of unknown boundary impedance, and showing that a similar biconvex formulation and optimization could solve this new problem efficiently (conference paper accepted for publication in ICASSP 2016 [38], see also Section 7.3.2).

7.2. Activities on waveform design for telecommunications

Peak to Average Power Ratio (PAPR), Orthogonal Frequency Division Multiplexing (OFDM), Generalized Waveforms for Multi Carrier (GWMC)

7.2.1. Characterizing multi-carrier waveform systems with optimum PAPR

Participant: Rémi Gribonval.

Main collaboration: Marwa Chafii, Jacques Palicot, Carlos Bader (Equipe SCEE, Supelec, Rennes)

In the context of the TEPN (Towards Energy Proportional Networks) Comin Labs project (see Section 9.1.1.2), in collaboration with the SCEE team at Supelec (thesis of Marwa Chafii co-supervised by R. Gribonval), we investigated a problem related to dictionary design: the characterization of waveforms with low Peak to Average Power Ratio (PAPR) for wireless communications. This is motivated by the importance of a low PAPR for energy-efficient transmission systems. A first stage of the work consisted in characterizing the statistical distribution of the PAPR for a general family of multi-carrier systems, leading to a journal paper [77] and several conference communications [75], [76]. The work this year concentrated on characterizing waveforms with optimum PAPR [30], [48].

7.3. Emerging activities on compressive learning and inverse problems

Compressive sensing, compressive learning, audio inpainting,

7.3.1. Audio inpainting

Participants: Rémi Gribonval, Nancy Bertin, Srđan Kitić.

Inpainting is a particular kind of inverse problems that has been extensively addressed in the recent years in the field of image processing.

Building upon our previous pioneering contributions (definition of the audio inpainting problem as a general framework for many audio processing tasks, application to the audio declipping or desaturation problem, formulation as a sparse recovery problem [59]), new results were obtained last year and this year to address the case of audio declipping with the competitive cospase approach. Last year, its promising results, especially when the clipping level is low, were confirmed experimentally by the formulation and use of a new algorithm named Cospase Iterative Hard Thresholding [95], which is a counterpart of the sparse Consistent Iterative Hard Thresholding.

This year, we proposed a new algorithmic framework called SPADE, based on non-convex heuristics and which can accommodate both the sparse and cospase prior. We studied their performance numerically and observed in particular that its cospase version offers a very appealing trade-off between reconstruction performance and computational time [31], making it suitable for practical applications, even in real-time. We could also confirm our results by subjective listening tests conducted this year [12].

The work on cospase audio declipping was awarded the Conexant best paper award at the LVA/ICA conference [31] and draw the attention of a world leading company in professional audio signal processing, with which some transfer has been negotiated.

Current and future works deal with developing advanced (co)sparse decomposition for audio inpainting, including several forms of structured sparsity (*e.g.* temporal and multichannel joint-sparsity), dictionary learning for inpainting, and several applicative scenarios (declipping, time-frequency inpainting, joint source separation and declipping).

7.3.2. *Blind Calibration of Impedance and Geometry*

Participants: Rémi Gribonval, Nancy Bertin, Srdan Kitic.

Main collaborations: Laurent Daudet, Thibault Nowakowski, Julien de Rosny (Institut Langevin)

This year, we also investigated extended inverse problem scenarios where a “lack of calibration” may occur, *i.e.*, when some physical parameters are needed for reconstruction but a priori unknown: speed of sound, impedance at the boundaries of the domain where the studied phenomenon propagates, or even the shape of these boundaries. In a first approach, based on our physics-driven cospase regularization of the sound source localization problem [23] (see section 7.1.3), we managed to preserve the sound source localization performance when the speed of sound is unknown, or, equally, when the impedance is unknown, provided the shape is and under some smoothness assumptions. Unlike the previous case (gain calibration), the arising problems are not convex but biconvex, and can be solved with proper biconvex formulation of ADMM algorithm. In a second approach based on eigenmode decomposition (limited to a 2D membrane), we showed that impedance learning with known shape, or shape learning with known impedance can be expressed as two facets of the same problem, and solved by the same approach, from a small number of measurements. Two papers presenting these two sets of results were accepted for publication in ICASSP 2016 [38], [35].

7.3.3. *Sketching for Large-Scale Mixture Estimation*

Participants: Rémi Gribonval, Nicolas Keriven.

Main collaborations: Patrick Perez (Technicolor R&I France) Anthony Bourrier (formerly Technicolor R&I France, now at GIPSA-Lab)

When fitting a probability model to voluminous data, memory and computational time can become prohibitive. In this work, we propose a framework aimed at fitting a mixture of isotropic Gaussians to data vectors by computing a low-dimensional sketch of the data. The sketch represents empirical moments of the underlying probability distribution. Deriving a reconstruction algorithm by analogy with compressive sensing, we experimentally show that it is possible to precisely estimate the mixture parameters provided that the sketch is large enough. Our algorithm provides good reconstruction and scales to higher dimensions than previous probability mixture estimation algorithms, while consuming less memory in the case of numerous data. It also provides a privacy-preserving data analysis tool, since the sketch does not disclose information about individual datum it is based on [70], [68], [69]. This year, we consolidated our extensions to non-isotropic Gaussians, with new

algorithms [49] and conducted large-scale experiments demonstrating its potential for speaker verification. A conference paper has been accepted to ICASSP 2016 [40] and a journal version is being finalized.

7.4. Recent results on tensor decompositions

tensor, multiway array, canonical polyadic decomposition, nonnegative tensor factorization

Multi-linear algebra is defined as the algebra of q -way arrays ($q > 2$), that is, the arrays whose elements are addressed by more than two indices. The first works dates back to Jordan who was interested in simultaneously diagonalizing two matrices at a time [93]. It is noteworthy that such two matrices can be interpreted as both slices of a three-way array and their joint diagonalization can be viewed as Hitchcock's polyadic decomposition [89] of the associated three-way array. Other works followed discussing rank problems related to multi-way structures and properties of multi-way arrays. However, these exercises in multilinear algebra were not linked to real data analysis but stayed within the realm of mathematics. Studying three-way data really started with Tucker's seminal work, which gave birth to the three-mode factor analysis [112]. His model is now often referred to as the Tucker3 model. At the same moment, other authors focused on a particular case of the Tucker3 model, calling it PARAFAC for PARAllel FACtor analysis [88], and on the means to achieve such a decomposition, which will become the famous canonical decomposition [73]. In honor to Hitchcock's pioneer work, we will call it the Canonical Polyadic (CP) decomposition.

Achieving a CP decomposition has been seen first as a mere non-linear least squares problem, with a simple objective criterion. In fact, the objective is a polynomial function of many variables, where some separate. One could think that this kind of objective is easy because smooth, and even infinitely differentiable. But it turns out that things are much more complicated than they may appear to be at first glance. Nevertheless, the Alternating Least Squares (ALS) algorithm has been mostly utilized to address this minimization problem, because of its programming simplicity. This should not hide the inherently complicated theory that lies behind the optimization problem. Moreover, in most of the applications, actual tensors may not exactly satisfy the expected model, so that the problem is eventually an approximation rather than an exact decomposition. This may result in a slow convergence (or lack of convergence) of iterative algorithms such as the ALS one [97]. Consequently, a new class of efficient algorithms able to take into account the properties of tensors to be decomposed is needed.

7.4.1. CP decomposition of semi-symmetric three-way arrays subject to arbitrary convex constraints

Participant: Laurent Albera.

Main collaborations : Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Jean-Christophe Pesquet (LIGM, France)

We addressed the problem of canonical polyadic decomposition of semi-symmetric 3rd order tensors (i.e. joint diagonalization by congruence) subject to arbitrary convex constraints. Sufficient conditions for the existence of a solution were proved. An efficient algorithm based on the Alternating Direction Method of Multipliers (ADMM) was then designed. ADMM provides an elegant approach for handling the additional constraint terms, while taking advantage of the structure of the objective function. Numerical tests on simulated matrices showed the benefits of the proposed method for low signal to noise ratios. Simulations in the context of nuclear magnetic resonance spectroscopy were also provided. This work was presented at the IEEE CAMSAP'15 conference [29].

7.4.2. Joint eigenvalue decomposition of non-defective matrices for the CP decomposition of tensors

Participant: Laurent Albera.

We proposed a fast and efficient Jacobi-like approach named JET (Joint Eigenvalue decomposition based on Triangular matrices) for the Joint EigenValue Decomposition (JEVD) of a set of real or complex non-defective matrices based on the LU factorization of the matrix of eigenvectors [98]. The JEVD can be useful in several contexts such as CP decomposition of tensors [99] and more particularly in Independent Component Analysis (ICA) based on higher order cumulants where it allows us to blindly compute the mixing matrix of sources with kurtosis of different signs. Regarding the proposed JET approach, contrary to classical Jacobi-like JEVD methods, its iterative procedure can be reduced to the search for only one of the two triangular matrices involved in the factorization of the matrix of eigenvectors, hence decreasing the numerical complexity. Two variants of the JET technique, namely JET-U and JET-O, which correspond to the optimization of two different cost functions were described in detail and these were extended to the complex case. Numerical simulations showed that in many practical cases the JET approach provides more accurate estimation of the matrix of eigenvectors than its competitors and that the lowest numerical complexity is consistently achieved by the JET-U algorithm.

7.5. Source separation and localization

Source separation, sparse representations, tensor decompositions, semi-nonnegative independent component analysis, probabilistic model, source localization

Source separation is the task of retrieving the source signals underlying a multichannel mixture signal.

About a decade ago, state-of-the-art approaches consisted of representing the signals in the time-frequency domain and estimating the source coefficients by sparse decomposition in that basis. These approaches rely only on spatial cues, which are often not sufficient to discriminate the sources unambiguously. Over the last years, we proposed a general probabilistic framework for the joint exploitation of spatial and spectral cues [106], which generalizes a number of existing techniques including our former study on spectral GMMs [61]. We showed how it could be used to quickly design new models adapted to the data at hand and estimate its parameters via the EM algorithm, and it became the basis of a large number of works in the field, including our own. In the last years, improvements were obtained through the use of prior knowledge about the source spatial covariance matrices [83], [92], [91], knowledge on the source positions and room characteristics [84], or a better initialization of parameters thanks to specific source localization techniques [67]. This accumulated progress lead to two main achievements last year: a new version of the Flexible Audio Source Separation Toolbox, fully reimplemented, was released [108] and we published an overview paper on recent and going research along the path of *guided* separation, *i.e.*, techniques and models allowing to incorporate knowledge in the process towards efficient and robust solutions to the audio source separation problem, in a special issue of IEEE Signal Processing Magazine devoted to source separation and its applications [113].

7.5.1. Towards real-world separation and remixing applications

Participants: Nancy Bertin, Frédéric Bimbot, Nathan Souviraà-Labastie, Ewen Camberlein, Romain Lebarbenchon.

Main collaboration: Emmanuel Vincent (EPI PAROLE, Inria Nancy)

While some challenges remain, work from previous years and our review paper on guided source separation [113] highlighted that progress has been made and that audio source separation is closer than ever to successful industrial applications, especially when some knowledge can be incorporated. This was exemplified by the contract with MAIA Studio, which reached its end in December 2014 and showed in particular how user input or side information could raise source separation tools to efficient solutions in real-world applications.

In some applicative contexts of source separation, several mixtures are available which contain similar instances of a given source. We have designed a general multi-channel source separation framework where additional audio references are available for one (or more) source(s) of a given mixture. Each audio reference is another mixture which is supposed to contain at least one source similar to one of the target sources. Deformations between the sources of interest and their references are modeled in a linear manner using a generic formulation. This is done by adding transformation matrices to an excitation-filter model, hence affecting different axes, namely frequency, dictionary component or time. A nonnegative matrix co-factorization algorithm

and a generalized expectation-maximization algorithm are used to estimate the parameters of the model. Different model parameterizations and different combinations of algorithms have been tested on music plus voice mixtures guided by music and/or voice references and on professionally-produced music recordings guided by cover references. Our algorithms has provided improvement to the signal-to-distortion ratio (SDR) of the sources with the lowest intensity by 9 to 15 decibels (dB) with respect to the original mixtures [25]. Combining these techniques, with automatic audio motif spotting, we have proposed a new concept called SPORES (for SPOtted Reference based Separation) and applied it to guided separation of audio tracks [13].

This year saw the beginning of a new industrial collaboration, in the context of the VoiceHome project, aiming at another challenging real-world application: natural language dialog in home applications, such as control of domotic and multimedia devices. As a very noisy and reverberant environment, home is a particularly challenging target for source separation, used here as a pre-processing for speech recognition (and possibly with stronger interactions with voice activity detection or speaker identification tasks as well). In 2015, we participated in a data collection campaign, and in benchmarking and adaptation of existing localization and separation tools to the particular context of this application.

7.5.2. *Implicit localization through audio-based control for robotics*

Participant: Nancy Bertin.

Main collaborations (audio-based control for robotics): Aly Magassouba and François Chaumette (Inria, EPI LAGADIC, France)

Acoustic source localization is, in general, the problem of determining the spatial coordinates of one or several sound sources based on microphone recordings. This problem arises in many different fields (speech and sound enhancement, speech recognition, acoustic tomography, robotics, aeroacoustics...) and its resolution, beyond an interest in itself, can also be the key preamble to efficient source separation. Common techniques, including beamforming, only provides the *direction of arrival* of the sound, estimated from the *Time Difference of Arrival (TDOA)* [67]. This year, we have particularly investigated alternative approaches, either where the explicit localization is not needed (audio-based control of a robot) or, on the contrary, where the exact location of the source is needed and/or TDOA is irrelevant (cospase modeling of the acoustic field, see Section 7.1.3).

In robotics, the use of aural perception has received recently a growing interest but still remains marginal in comparison to vision. Yet audio sensing is a valid alternative or complement to vision in robotics, for instance in homing tasks. Most existing works are based on the relative localization of a defined system with respect to a sound source, and the control scheme is generally designed separately from the localization system.

In contrast, the approach that we started investigating last year focuses on a sensor-based control approach. We proposed a new line of work, by considering the hearing sense as a direct and real-time input of closed loop control scheme for a robotic task. Thus, and unlike most previous works, this approach does not necessitate any explicit source localization: instead of solving the localization problem, we focus on developing an innovative modeling based on sound features. To address this objective, we placed ourselves in the sensor-based control framework, especially visual servoing (VS) that has been widely studied in the past [78].

From now on, we have established an analytical model linking sound features and control input of the robot, defined and analyzed robotic homing tasks involving multiple sound sources, and validated the proposed approach by simulations and experiments with an actual robot. This work is mainly lead by Aly Magassouba, whose Ph.D. is co-supervised by Nancy Bertin and François Chaumette. A conference paper presenting these first results was published this year [34] and another was submitted to ICRA 2016. Future work will include additional real-world experiments with the robot Romeo from Aldebaran Robotics, investigation of new tasks with active sensing strategies, explicit use of echoes and reverberation to increase robustness, and exploration of dense methods (control from raw acoustic signals rather than from acoustic features).

7.5.3. *Brain source localization*

Participants: Laurent Albera, Srđan Kitić, Nancy Bertin, Rémi Gribonval.

Main collaborations : Hanna Becker (GIPSA & LTSI, France), Pierre Comon (GIPSA, France), Isabelle Merlet (LTSI, France), Fabrice Wendling (LTSI, France)

From tensor to sparse models

The brain source imaging problem has been widely studied during the last decades, giving rise to an impressive number of methods using different priors. Nevertheless, a thorough study of the latter, including especially sparse and tensor-based approaches, is still missing. Consequently, we proposed i) a taxonomy of the methods based on a priori assumptions, ii) a detailed description of representative algorithms, iii) a review of identifiability results and convergence properties of different techniques, and iv) a performance comparison of the selected methods on identical data sets. Our aim was to provide a reference study in the biomedical engineering domain which may also be of interest for other areas such as wireless communications, audio source localization, and image processing where ill-posed linear inverse problems are encountered and to identify promising directions for future research in this area. This work was published in the IEEE Signal Processing Magazine [14].

A sparsity-based approach

Identifying the location and spatial extent of several highly correlated and simultaneously active brain sources from EEG recordings and extracting the corresponding brain signals is a challenging problem. In our comparison of source imaging techniques presented at ICASSP'14 [65], the VB-SCCD algorithm [81], which exploits the sparsity of the variational map of the sources, proved to be a promising approach. We proposed several ways to improve this method. In order to adjust the size of the estimated sources, we added a regularization term that imposes sparsity in the original source domain. Furthermore, we demonstrated the application of ADMM, which permitted to efficiently solve the optimization problem. Finally, we also considered the exploitation of the temporal structure of the data by employing L1,2-norm regularization. The performance of the resulting algorithm, called Sissy, was evaluated based on realistic simulations in comparison to VB-SCCD and several state-of-the-art techniques for extended source localization. This work was partially presented at EUSIPCO'14 [64] and a journal paper is in preparation.

Tensor- and sparsity-based approaches

The separation of EEG sources is a typical application of tensor decompositions in biomedical engineering. The objective of most approaches studied in the literature consists in providing separate spatial maps and time signatures for the identified sources. However, for some applications, a precise localization of each source is required.

To achieve this, a two-step approach was presented at the IEEE EMBC conference [26]. The idea of this approach is to separate the sources using the canonical polyadic decomposition in the first step and to employ the results of the tensor decomposition to estimate distributed sources in the second step, using the Sissy algorithm [64].

Next, we proposed to combine the tensor decomposition and the source localization in a single step [27]. To this end, we directly imposed structural constraints, which are based on a priori information on the possible source locations, on the factor matrix of spatial characteristics. The resulting optimization problem was solved using the alternating direction method of multipliers (ADMM), which was incorporated in the alternating least squares tensor decomposition algorithm. Realistic simulations with epileptic EEG data confirmed that the proposed single-step source localization approach outperformed the previously developed two-step approach.

7.5.4. Independent component analysis

Participant: Laurent Albera.

Main collaboration: Sepideh Hajipour (LTSI & BiSIPL), Isabelle Merlet (LTSI, France), Mohammad Bagher Shamsollahi (BiSIPL, Iran)

Independent Component Analysis (ICA) is a very useful tool to process biomedical signals including EEG data.

We proposed a Jacobi-like Deflationary ICA algorithm, named JDICA. More particularly, while a projection-based deflation scheme inspired by Delfosse and Loubaton's ICA technique (DeLL[®]) [80] was used, a Jacobi-like optimization strategy was proposed in order to maximize a fourth order cumulant-based contrast built from whitened observations. Experimental results obtained from simulated epileptic data mixed with a real muscular activity and from the comparison in terms of performance and numerical complexity with the FastICA [90], RobustICA [114] and DeLL[®] algorithms, show that the proposed algorithm offers the best trade-off between performance and numerical complexity. This work was published in the IEEE Signal Processing Letters journal [21].

In addition, we illustrated in the ICA context the interest of being able to solve efficiently the (non-orthogonal) JEVD problem. More particularly, we showed that, when the noise covariance matrix is unknown and the source kurtoses have different signs, the joint diagonalization problem involved in the ICAR method [58] becomes a non-orthogonal JEVD problem. Consequently, by using our JET-U algorithm [98], giving birth to the MICAR-U (Modified ICAR based on JET-U) technique, we then provided a more robust ICA method. The identifiability of the MICAR-U technique was studied and proved under some conditions. Computer results given in the context of brain interfaces showed the better ability of the MICAR-U approach to denoise electrocortical data compared to classical ICA techniques for low signal to noise ratio values. These results were presented in [98].

7.5.5. *Semi-nonnegative independent component analysis*

Participant: Laurent Albera.

Main collaboration: Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Huazhong Shu (LIST, China)

ICA plays also an important role in many other areas including speech and audio [62], [63], [74], [71], radiocommunications [79] and document restoration [111] to cite a few.

For instance in [111], the authors use ICA to restore digital document images in order to improve the text legibility. Indeed, under the statistical independence assumption, authors succeed in separating foreground text and bleed-through/show-through in palimpsest images. Furthermore, authors in [82] use ICA to solve the ambiguity in X-ray images due to multi-object overlappings. They presented a novel object decomposition technique based on multi-energy plane radiographs. This technique selectively enhances an object that is characterized by a specific chemical composition ratio of basis materials while suppressing the other overlapping objects. Besides, in the context of classification of tissues and more particularly of brain tumors [107], ICA is very effective. In fact, it allows for feature extraction from Magnetic Resonance Spectroscopy (MRS) signals, representing them as a linear combination of tissue spectra, which are as independent as possible [110]. Moreover, using the JADE algorithm [72] applied to a mixture of sound waves computed by means of the constant-Q transform (Fourier transform with log-frequency) of a temporal waveform broken up into a set of time segments, the authors of [71] describe trills as a set of note pairs described by their spectra and corresponding time envelopes. In this case, pitch and timing of each note present in the trill can be easily deduced.

All the aforementioned applications show the high efficiency of the ICA and its robustness to the presence of noise. Despite this high efficiency in resolving the proposed applicative problems, authors did not fully exploit properties enjoyed by the mixing matrix such as its nonnegativity. For instance in [82], the thickness of each organ, which stands for the mixing coefficient, is real positive. Furthermore, reflectance indices in [111] for the background, the overwriting and the underwriting, which correspond to the mixing coefficients, are also nonnegative. Regarding tissue classification from MRS data, each observation is a linear combination of independent spectra with positive weights representing concentrations [87]; the mixing matrix is again nonnegative.

By imposing the nonnegativity of the mixing matrix within the ICA process, we showed through computer results that the extraction quality can be improved. Exploiting the nonnegativity property of the mixing matrix during the ICA process gives rise to what we call semi-nonnegative ICA. More particularly, we performed the latter by computing a constrained joint CP decomposition of cumulant arrays of different orders [100]

having the nonnegative mixing matrix as loading matrices. After merging the entries of the cumulant arrays in the same third order array, the reformulated problem follows the semi-symmetric semi-nonnegative CP model defined in section 7.4.1. Hence we use the new method described in section 7.4.1 to perform semi-nonnegative ICA. Performance results in biomedical engineering were given in the paper cited in section 7.4.1.

7.6. Audio and speech content processing

Audio segmentation, speech recognition, motif discovery, audio mining

7.6.1. Audio motif discovery and spotting

Participants: Frédéric Bimbot, Nathan Souviraà-Labastie.

This work was performed in close collaboration with Emmanuel Vincent from Inria Nancy-Grand Est.

As an alternative to supervised approaches for multimedia content analysis, where predefined concepts are searched for in the data, we investigate content discovery approaches where knowledge emerge from the data. Following this general philosophy, we pursued work on motif discovery in audio contents.

Audio motif discovery is the task of finding out, without any prior knowledge, all pieces of signals that repeat, eventually allowing variability. The developed algorithms allows discovering and collecting occurrences of repeating patterns in the absence of prior acoustic and linguistic knowledge, or training material. When the audio pattern is determined in a user supervised fashion, the task becomes that of motif spotting.

Investigated in the context of SPORES (SPotted Reference based Separation) [13], audio motif spotting has been illustrated as a useful way to exploit redundancy in audio contents, for guided source separation purposes.

7.6.2. Mobile device for the assistance of users in potentially dangerous situations

Participants: Romain Lebarbenchon, Ewen Camberlein, Frédéric Bimbot.

The S-Pod project is a cooperative project between industry and academia aiming at the development of mobile systems for the detection of potentially dangerous situations in the immediate environment of a user, without requiring his/her active intervention.

In this context, the PANAMA research group has been involved in the design of algorithms for the analysis and monitoring of the acoustic scene around the user, yielding audio-based information which can be fused with other sensors (physiological, positional, etc.) in order to trigger an alarm (and subsequent appropriate measures) when needed.

The last phase of the project has been dedicated towards robustness improvement of audio scene analysis, with a particular focus on threat vs non-threat detection, on the basis of adaptive training scenarii. Knowledge and know-how transfer has been achieved for the hardware implementation of the designed methods and the efficient integration into an operational prototype.

7.7. Music Content Processing and Music Information Retrieval

Acoustic modeling, non-negative matrix factorisation, music language modeling, music structure

7.7.1. Music structure modeling by System & Contrast

Participants: Frédéric Bimbot, Corentin Louboutin.

The *System & Contrast* (S&C) model aims at describing the inner organization of structural segments within music pieces in terms of : (i) a carrier system, i.e. a sequence of morphological elements forming a multi-dimensional network of self-deducible syntagmatic relationships and (ii) a contrast, i.e. a substitutive element, usually the last one, which partly departs from the logic implied by the rest of the system [16].

With a primary focus on pop music, the S&C model provides a framework to describe internal implication patterns in musical segments by encoding similarities and relations between its constitutive elements so as to minimize the complexity of the resulting description. It is applicable at several timescales and to a wide variety of musical dimensions in a polymorphous way, therefore offering an attractive meta-description of different types of musical contents.

We have established the filiation of the S&C model as an extension of Narmour's Implication-Realization model [104], [105] and Cognitive Rule-Mapping [103].

We have introduced the Minimum Description Length scheme as a productive paradigm that supports the estimation of S&C descriptions and establishes promising connections between Music Data Processing and Information Retrieval on the one hand, and modern theories in Music Perception and Cognition on the other hand, together with interesting perspectives in other areas in Musicology.

The model is currently being investigated for the multi-scale description of chord sequences.

7.7.2. *Tree-based representation of music pieces*

Participants: Frédéric Bimbot, Corentin Guichaoua.

Modeling music structure, i.e. the organisation of musical elements and their relationships within a piece of music, is an open problem of primary importance in MIR.

To address this challenge, we approach music structure description as the inference of a low complexity generative grammar able to account for the music piece, itself represented as a sequence of symbols.

Originally introduced for the inference of structure in DNA sequences, Straight-Line Grammars (SLG) form a particular subclass of Context-Free Grammars (CFG) which can be used to model symbolic sequences and to represent them as hierarchical trees. However, SLGs appear to be poorly suited to some particularities of musical patterns, such as segmental regularities, closure substitutions and specific style structures.

We are designing and investigating formal and algorithmic extensions of SLGs as SLEGs (Straight-Line Edition Grammars). Based on a more general minimum description criterion, the SLEG extension allows alterations in the generation step and enables the use of priors in the grammar inference process. Current work includes a diagnostic comparison between the various approaches on the structural segmentation of chord sequences from pop songs.

PERCEPTION Project-Team

6. New Results

6.1. Supervised Audio-Source Localization

We addressed the problem of localizing audio sources using binaural measurements. After proposing an unsupervised method [20], we proposed a supervised formulation that simultaneously localizes multiple sources at different locations [22]. The approach is intrinsically efficient because, contrary to prior work, it relies neither on source separation, nor on monaural segregation. The method starts with a training stage that establishes a locally-linear Gaussian regression [21] between the directional coordinates of all the sources and the auditory features extracted from binaural measurements. While fixed-length wide-spectrum sounds (white noise) are used for training to reliably estimate the model parameters, we show that the testing (localization) can be extended to variable-length sparse-spectrum sounds (such as speech), thus enabling a wide range of realistic applications. Indeed, we demonstrate that the method can be used for audio-visual fusion, namely to map speech signals onto images and hence to spatially align the audio and visual modalities, thus enabling to discriminate between speaking and non-speaking faces. We release a novel corpus of real-room recordings that allow quantitative evaluation of the co-localization method in the presence of one or two sound sources. Experiments demonstrate increased accuracy and speed relative to several state-of-the-art methods. More recently the method has been extended to an arbitrary number of microphones [35], [34]. Moreover, we have started to develop a method that extracts the direct path on an acoustic wave in order to enable robust audio-source localization in reverberant environments [40].

Websites:

<https://team.inria.fr/perception/research/acoustic-learning/>

<https://team.inria.fr/perception/research/binaural-ssl/>

<https://team.inria.fr/perception/research/local-rtf/>

6.2. Multichannel Audio-Source Separation

We address the problem of separating audio sources from time-varying convolutive mixtures. We proposed an unsupervised probabilistic framework based on the local complex-Gaussian model combined with non-negative matrix factorization. The time-varying mixing filters are modeled by a continuous temporal stochastic process. This model extends the case of static filters which corresponds to static audio sources. While static filters can be learned in advance, e.g. [37], time-varying filters cannot and therefore the problem is more complex. We present a variational expectation-maximization (VEM) algorithm that employs a Kalman smoother to estimate the time-varying mixing matrix, and that jointly estimates the source parameters. The sound sources are then separated by Wiener filters constructed with the estimators provided by the VEM algorithm. Extensive experiments on simulated data show that the proposed method outperforms a block-wise version of a state-of-the-art baseline method. This work is part of the PhD topic of Dionyssos Kounades Bastian and is conducted in collaboration with Sharon Gannot (Bar Ilan University) and Xavier Alameda Pineda (University of Trento). It received the best student paper award at WASPAA'15 [31]. An extended version has been submitted to IEEE Transactions on Audio, Speech, and Language Processing [39].

6.3. Audio-Visual Speaker Tracking and Recognition

Any multi-party conversation system benefits from speaker diarization, that is, the assignment of speech signals among the participants. More generally, in HRI and CHI scenarios it is important to recognize the speaker over time. We propose to address speaker diarization and speaker recognition using both audio and visual data. We cast the diarization problem into a tracking formulation whereby the active speaker is detected and tracked over time. A probabilistic tracker exploits the spatial coincidence of visual and auditory observations and infers a single latent variable which represents the identity of the active speaker. Visual and auditory observations are fused using our recently developed weighted-data mixture model [38], while several options for the speaking turns dynamics are fulfilled by a multi-case transition model. The modules that translate raw audio and visual data into image observations are also described in detail. The performance of the proposed trackers [29], [30] are tested on challenging data-sets that are available from recent contributions which are used as baselines for comparison. Currently we are developing a variational framework for the on-line tracking of multiple persons [36].

Websites:

<https://team.inria.fr/perception/research/speakerloc/>

<https://team.inria.fr/perception/research/speechockdet/>

<https://team.inria.fr/perception/research/avdiarization/>

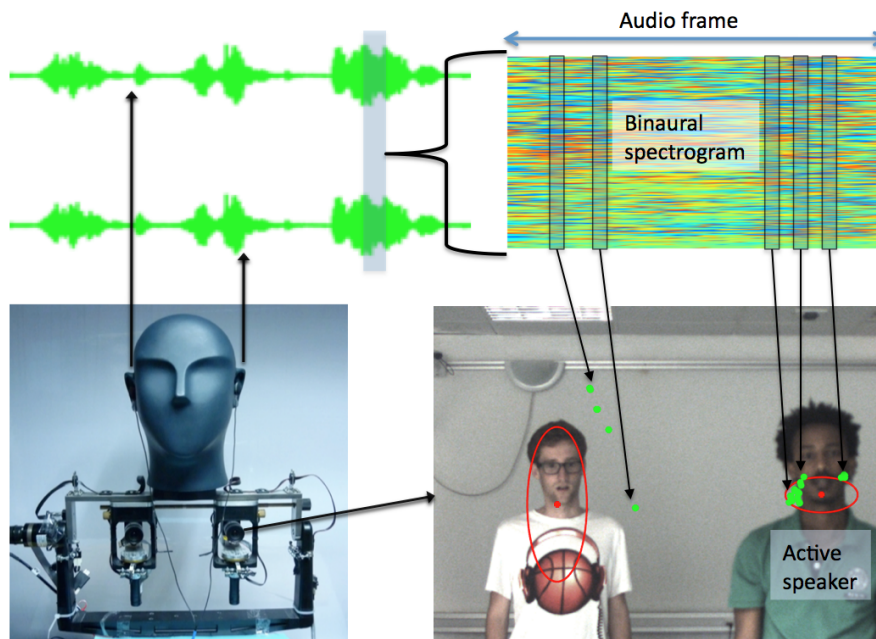


Figure 6. This figure illustrates the general principle of our audio-visual speaker tracking and diarization method. The auditory and visual data are recorded with two microphones and one camera. The audio signals are segmented into frames and each frame (vertical grey rectangle) is transformed into a binaural spectrogram [20]. This spectrogram is composed of a sequence of binaural vectors (vertical rectangles) and each binaural vector is mapped onto a sound-source direction which corresponds to a point in the image plane (green dots) [22]. The proposed audio-visual tracker associates people detected in the image sequence with these sound directions via audio-visual clustering [38] that is combined with an active-speaker transition model.

6.4. Head Pose Estimation

Head pose estimation is an important task, because it provides information about cognitive interactions that are likely to occur. Estimating the head pose is intimately linked to face detection. We addressed the problem of head pose estimation with three degrees of freedom (pitch, yaw, roll) from a single image and in the presence of face detection errors. Pose estimation is formulated as a high-dimensional to low-dimensional mixture of linear regression problem [21]. We propose a method that maps HOG-based descriptors, extracted from face bounding boxes, to corresponding head poses. To account for errors in the observed bounding-box position, we learn regression parameters such that a HOG descriptor is mapped onto the union of a head pose and an offset, such that the latter optimally shifts the bounding box towards the actual position of the face in the image. The performance of the proposed method is assessed on publicly available datasets. The experiments that we carried out show that a relatively small number of locally-linear regression functions is sufficient to deal with the non-linear mapping problem at hand. Comparisons with state-of-the-art methods show that our method outperforms several other techniques. This work is part of the PhD of Vincent Drouard and it received the best student paper award (second place) at the IEEE ICIP'15 [28]. Currently we investigate a temporal extension of this model.

Website:

<https://team.inria.fr/perception/research/head-pose/>

6.5. High-Resolution Scene Reconstruction

We addressed the problem of range-stereo fusion for the construction of high-resolution depth maps. In particular, we combine low-resolution depth data with high-resolution stereo data, in a maximum a posteriori (MAP) formulation. Unlike existing schemes that build on MRF optimizers, we infer the disparity map from a series of local energy minimization problems that are solved hierarchically, by growing sparse initial disparities obtained from the depth data. The accuracy of the method is not compromised, owing to three properties of the data-term in the energy function. Firstly, it incorporates a new correlation function that is capable of providing refined correlations and disparities, via sub-pixel correction. Secondly, the correlation scores rely on an adaptive cost aggregation step, based on the depth data. Thirdly, the stereo and depth likelihoods are adaptively fused, based on the scene texture and camera geometry. These properties lead to a more selective growing process which, unlike previous seed-growing methods, avoids the tendency to propagate incorrect disparities. The proposed method gives rise to an intrinsically efficient algorithm, which runs at 3FPS on 2.0MP images on a standard desktop computer. The strong performance of the new method is established both by quantitative comparisons with state-of-the-art methods, and by qualitative comparisons using real depth-stereo data-sets [23]. This work is funded by the ANR project MIXCAM.

Website:

<https://team.inria.fr/perception/research/dsfusion/>

6.6. Hyper-Spectral Image Analysis

As an extension to our work on high-dimensional regression [21] we addressed the problem of analyzing hyper-spectral data. In particular we addressed the problem of recovering physical properties (parameters) from hyper-spectral low-resolution images, i.e. at large planetary scales. This involves resolving inverse problems which can be addressed within machine learning, with the advantage that, once a relationship between physical parameters and spectra has been established in a data-driven fashion, the learned relationship can be used to estimate physical parameters for new hyper-spectral observations. Within this framework, we propose a spatially-constrained and partially-latent regression method which maps high-dimensional inputs (hyper-spectral images) onto low-dimensional responses (physical parameters such as the local chemical composition of the soil). The proposed regression model comprises two key features. Firstly, it combines a Gaussian mixture of locally-linear mappings (GLLiM) with a partially-latent response model. While the former makes high-dimensional regression tractable, the latter enables to deal with physical parameters that cannot be observed or, more generally, with data contaminated by experimental artifacts that cannot be

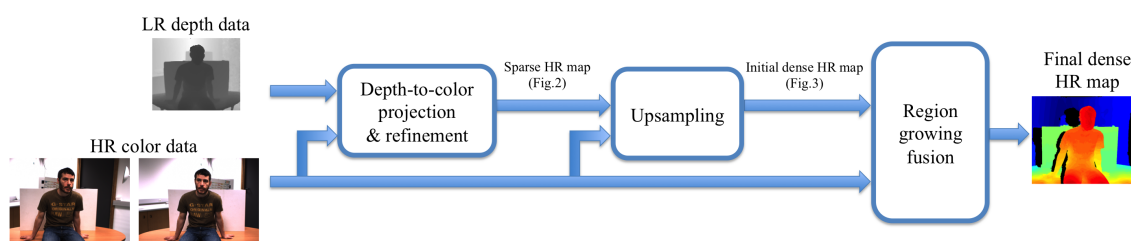


Figure 7. The pipeline of the proposed depth-stereo fusion method [23]. The low-resolution (LR) depth data are projected onto the color data and refined to yield a high-resolution (HR) sparse disparity map. Starting from these disparity seeds, an upsampling process provides an initial HR dense disparity map. Both the HR seeds and the initial dense disparity map are then used by the region-growing depth-stereo fusion to produce the final HR depth map. A prominent feature of our method is that fusion takes place at several data processing stages.

explained with noise models. Secondly, spatial constraints are introduced in the model through a Markov random field (MRF) prior which provides a spatial structure to the Gaussian-mixture hidden variables [19]. Experiments conducted on a database composed of remotely sensed observations collected from the Mars planet by the Mars Express orbiter demonstrate the effectiveness of the proposed model.

6.7. Gaussian Mixture Regression for Acoustic-Articulatory Inversion

The team expertise in latent-variable mixture models was applied to the problem of adaptation of an acoustic-articulatory model of a reference speaker to the voice of another speaker, using a limited amount of audio-only data [25]. In the context of pronunciation training, a virtual talking head displaying the internal speech articulators (e.g., the tongue) could be automatically animated by means of such a model using only the speaker's voice. In this study, the articulatory-acoustic relationship of the reference speaker is modeled by a gaussian mixture model (GMM). To address the speaker adaptation problem, we propose a new framework called cascaded Gaussian mixture regression (C-GMR), and derive two implementations. The first one, referred to as Split-C-GMR, is a straightforward chaining of two distinct GMRs: one mapping the acoustic features of the source speaker into the acoustic space of the reference speaker, and the other estimating the articulatory trajectories with the reference model. In the second implementation, referred to as Integrated-C-GMR, the two mapping steps are tied together in a single probabilistic model. For this latter model, we present the full derivation of the exact EM training algorithm, that explicitly exploits the missing data methodology of machine learning. Other adaptation schemes based on maximum-a posteriori (MAP), maximum likelihood linear regression (MLLR) and direct cross-speaker acoustic-to-articulatory GMR are also investigated. Experiments conducted on two speakers for different amount of adaptation data show the interest of the proposed C-GMR techniques. This work was done in collaboration with Thomas Hueber and Gérard Bailly from Gipsa Lab and with Xavier Alameda-Pineda from University of Trento and former team member.

POTIOC Project-Team

7. New Results

7.1. Pointing in Spatial Augmented Reality from 2D Pointing Devices

Participants: Renaud Gervais, Jérémy Frey, Martin Hachet.



Figure 3. A user reaches a target displayed on a spatially augmented object with an indirect input device

Spatial Augmented Reality (SAR) opens interesting perspectives for new generations of mixed reality applications. Compared to traditional human-computer interaction contexts, there is little work that studies user performance in SAR. In this project, we present an experiment that compares pointing in SAR versus pointing in front of a screen, from standard pointing devices (mouse and graphics tablet). The results showed that the participants tend to interact in SAR in a way that is similar to the screen condition, without a big loss of performance [30] (See Figure 3).

7.2. Tangible Viewports

Participants: Renaud Gervais, Joan Sol Roo, Martin Hachet.

Spatial augmented reality and tangible interaction enrich the standard computer I/O space. Systems based on such modalities offer new user experiences and open up interesting perspectives in various fields. On the other hand, such systems tend to live outside the standard desktop paradigm and, as a consequence, they do not benefit from the richness and versatility of desktop environments. In this work, we propose to join together physical visualization and tangible interaction within a standard desktop environment. We introduce the concept of Tangible Viewport, an on-screen window that creates a dynamic link between augmented objects and computer screens, allowing a screen-based cursor to move onto the object in a seamless manner (Figure 4). We describe an implementation of this concept and explore the interaction space around it. A preliminary evaluation shows that the metaphor is transparent to the users while providing the benefits of tangibility [31].

7.3. Tobe

Participants: Renaud Gervais, Jérémy Frey, Alexis Gay, Fabien Lotte, Martin Hachet.

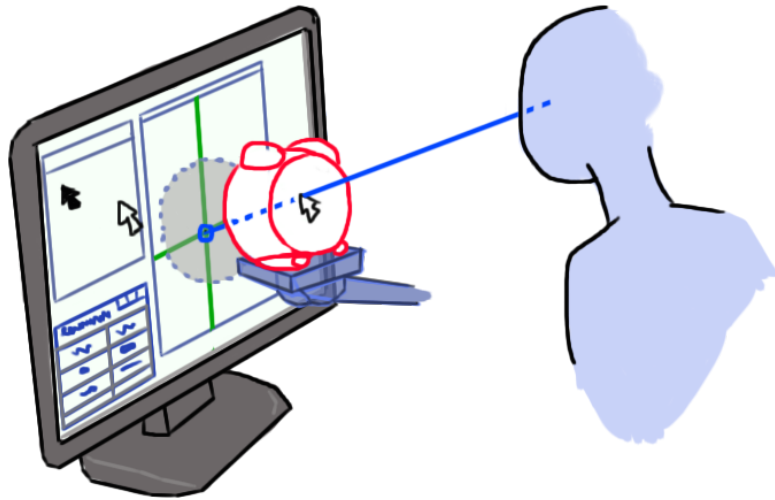


Figure 4. A user interacts with an object located in front of the screen as if the object was rendered on screen

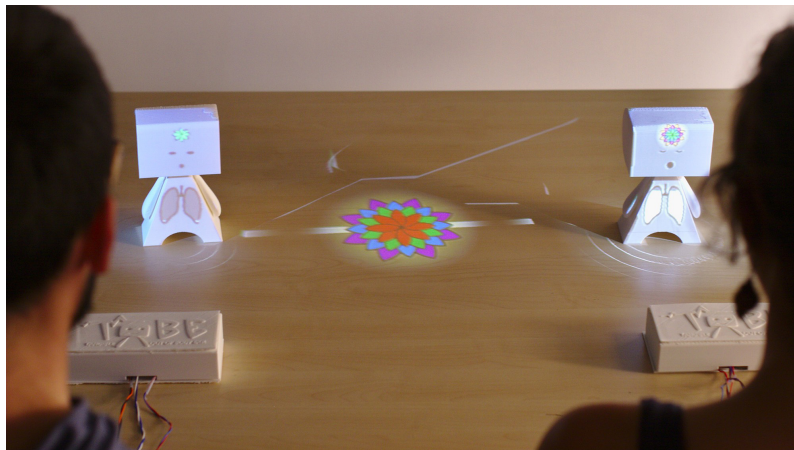


Figure 5. Two users are relaxing together using Tobe as a biofeedback for heartrate and breathing

We propose Tobe, a toolkit for creating Tangible Out-of-Body Experiences: exposing the inner states of users using physiological signals such as heart rate or brain activity. Tobe can take the form of a tangible avatar displaying live physiological readings to reflect on ourselves and others. Such a toolkit could be used by researchers and designers to create a multitude of potential tangible applications, including (but not limited to) educational tools about Science Technologies Engineering and Mathematics (STEM) and cognitive science, medical applications or entertainment and social experiences with one or several users or Tobes involved. Through a co-design approach, we investigated how everyday people picture their physiology and we validated the acceptability of Tobe in a scientific museum. We also give a practical example where two users relax together, with insights on how Tobe helped them to synchronize their signals and share a moment, as illustrated in Figure 5 [29].

7.4. Inner Garden

Participants: Joan Sol Roo, Renaud Gervais, Martin Hachet.

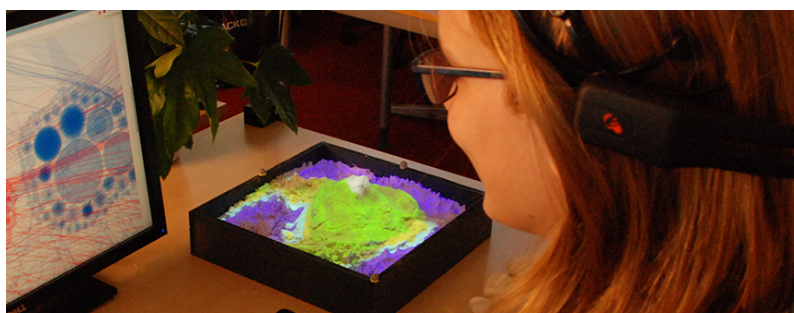


Figure 6. Inner garden is an augmented sandbox which depicts an evolving world reflecting the inner state of the user

We present a prototype of an augmented sandbox where the sand is used to create a miniature living world, designed as an ambient display for contemplation and self-reflection. The landscape can be reshaped at any time. Once the sand is left still for a moment, the world starts evolving – vegetation grows, water flows and creatures move around – according to the user’s internal state. We use a consumer-grade EEG and breathing sensors to reflect on frustration and meditative states of users, which they can monitor by looking at the sandbox (Figure 6) [49].

7.5. Augmented geographic maps

Participants: Julia Chatain, Marie Demangeat, Anke Brock, Martin Hachet.

Interactive geographic maps are today widely available, but remain mostly limited to standard interaction contexts. We introduce SyMAPse [48], a spatial augmented reality map, based on the **PapARt framework**. In our prototype, we use augmented reality to display a virtual map on a physical piece of paper, thus keeping features of both media. Thanks to the digital map base, users can pan, zoom and even change the basemap. At the same time, the paper base allows users to manipulate the map physically and so to interact in a more “natural” way, as well as to draw on the paper using regular pens. In a preliminary study with visitors of the "Cap Sciences" science center, we compared interaction techniques based on touch, tangible and spatial modalities for these three common map functions: zooming, panning, and changing the basemap. Our results suggest that object-based and spatial interaction may be advantageous over touch in our augmented reality setup.



Figure 7. Interacting with an augmented geographic map using tangible, spatial and multitouch interaction

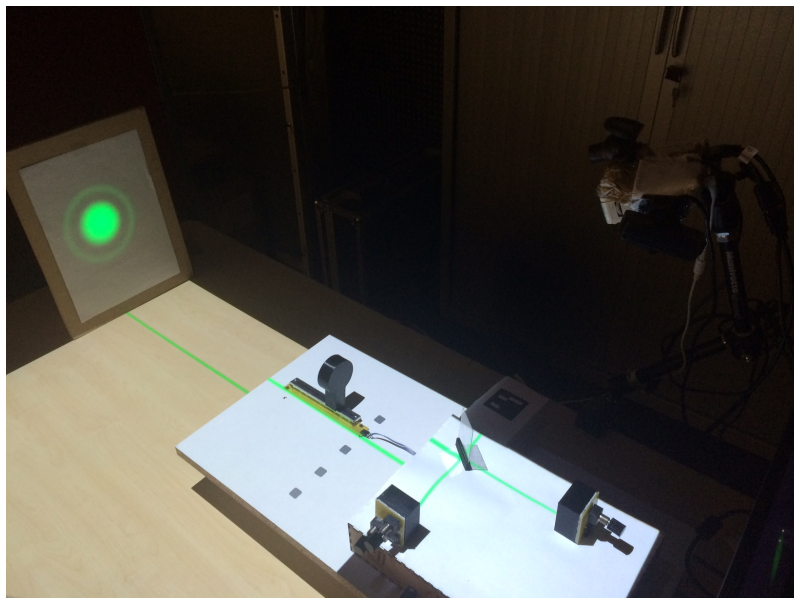


Figure 8. HOBIT mixes physical and virtual elements to teach optics.

7.6. HOBIT - Hybrid Optical Bench for Innovative Teaching

Participants: David Furio, Benoit Coulais, Martin Hachet.

Experiments in optics are essential for learning and understanding physical phenomena. The problem with these experiments is that they are generally time consuming for both their construction and their maintenance, potentially dangerous through the use of laser sources, and often expensive due to high technology optical components. We propose to simulate such experiments by using hybrid systems that exploit both spatial augmented reality and tangible interaction (See Figure 8). In particular, we focus on one of the most popular optical experiments: Michelson interferometer. In our approach, we target a highly interactive system where students are able to interact in real time with the Augmented Michelson Interferometer (AMI) to observe, test hypotheses and then to enhance their comprehension. Compared to a fully digital simulation, we are investigating an approach that benefits from both physical and virtual elements, and where the students experiment by manipulating 3D-printed physical replicas of optical components (e.g. lenses and mirrors). Our objective is twofold. First, we want to ensure that the students will learn with our simulator the same concepts and skills that they learn with traditional methods. Second, we hypothesize that such a system opens new opportunities to teach optics in a way that was not possible before, by manipulating concepts beyond the limits of observable physical phenomena. To reach this goal, we have built a complementary team composed of experts in the field of optics, human-computer interaction, computer graphics, sensors and actuators, and education science. HOBIT is a joint project between Inria and Université de Bordeaux (IDEX CPU – LAPHIA), in collaboration with Université de Lorraine (team PErSEUs). [28]

7.7. Mixed Reality to improve children's interaction with astronomical concepts

Participant: Martin Hachet.

This project stands on a collaboration with Stéphanie Fleck from Université de Lorraine. To make astronomical learning more efficient for young pupils, we have designed an Augmented Inquiry-Based Learning Environment (AIBLE): HELIOS. Because manipulations in astronomy are intrinsically not possible, we propose to manipulate the underlying model. With HELIOS, virtual replicas of the sun, moon and earth are directly manipulated from tangible manipulations. This digital support combines the possibilities of Augmented Reality (AR) while maintaining intuitive interactions following the principles of didactic of sciences. Light properties are taken into account and shadows of Earth and Moon are directly produced by an omnidirectional light source associated to the virtual Sun. This AR environment provides users with experiences they would otherwise not be able to experiment in the physical world. Our main goal is that students can take active control of their learning, express and support their ideas, make predictions and hypotheses, and test them by conducting investigations. [24][23]

7.8. Combining and Revealing Spaces for Musical Performances

Participant: Martin Hachet.

In collaboration with University of Bristol (Florent Berthaut, Diego Martinez, and Sriram Subramanian) we have designed a mixed-reality environment for musical performances that allows for freely displaying virtual content on stage, such as 3D virtual musical interfaces or visual augmentations of instruments and performers. This environment, called Reflets, relies on spectators and performers revealing virtual objects by slicing through them with body parts or objects, and on planar slightly reflective transparent panels that combine the stage and audience spaces. It allows for placing virtual content anywhere on large stages, even overlapping with physical elements and provides a consistent rendering of this content for large numbers of spectators. It also preserves non-verbal communication between the audience and the performers, and is inherently engaging for the spectators. Reflets opens musical performance opportunities such as augmented interaction between musicians and novel techniques for 3D sound shapes manipulation [20].

7.9. Improving User-Training for Brain-Computer Interfaces

Participants: Martin Hachet, Emilie Jahanpour, Camille Jeunet, Fabien Lotte, Boris Masencal, Julia Schumacher.

While Mental Imagery based Brain-Computer Interfaces (MI-BCIs) are promising for many applications, ranging from assistive technologies for motor disabled patients to video games, their usability “out-of-the-lab” has been questioned due to their lack of reliability: literature reports that 15% to 30% of users cannot control such a technology, while most of the remaining users obtain only modest performances. As controlling an MI-BCI requires the acquisition of specific skills (i.e., producing stable and distinct brain-activity patterns), an adapted training is necessary. Thus, the main objective of our project is to improve the user training to facilitate the acquisition of MI-BCI related skills. In order to do so, we focused on two axes [18]: (1) the impact of the user-profile and (2) the impact of the protocol on MI-BCI performance.

Concerning the impact of the user-profile, our results ([40], [14]) suggested an important impact of some aspects of the personality (such as the tension and autonomy levels) as the spatial abilities (i.e., the ability to produce, interpret and transform mental imageries). On the one hand, we are working on learning companions, whose goal would be to provide the learners with a specific emotional support, based on their profile and on their cognitive state. On the other hand, we are currently implementing and testing a spatial ability training in order to test the hypothesis of a causal effect of the spatial abilities on MI-BCI performance [39]. In other words, we would like to know if increasing spatial abilities would result in better MI-BCI performance. One application of such a research is stroke rehabilitation. Indeed, motor after-effects are usual following a stroke. MI-BCI have been shown very useful to facilitate the rehabilitation process, which consists in enhancing brain plasticity through motor-imagery, as they enable to visualise the BCI activity while the patients perform MI-tasks. However, MI-tasks tend to increase the depressive state of the patients as they remind them they lost the ability to move their limb. Thus, as spatial ability exercises (e.g., mental rotation) activate the motor cortex, they could be used as more transparent rehabilitation exercises to trigger brain plasticity.

Second, concerning the impact of the protocol, we completed a study (see activity report 2014) in which we asked the participants to use the standard MI-BCI training protocol to learn to perform simple motor tasks: drawing circles and triangles on a graphic tablet. As it would have been the case for an MI-BCI experiment, they had to find the right strategy so that the system recognises the task they were performing. Seventeen percent of the participants (N=54) showed difficulties in performing these tasks. Also, when we selected the 10 best and 10 worst performers of this experiment and asked them to use an MI-BCI (by imagining left and right-hand movements), it appeared that the ones who had difficulties in performing the simple motor tasks improved in terms of performance during the MI-BCI experiment, while the participants who performed well during the motor experiment did not progress during the second. Furthermore, we have shown that tactile feedback was more efficient than an equivalent visual feedback in a multitasking context [32]. Based on a literature review, this could be due to an increased sense of agency (i.e., the feeling to be in control). We are thus currently exploring the impact of the sense of agency on MI-BCI performance. Finally, still regarding the feedback, we explored what kind of information could help the user to perform better mental imagery tasks. As such, we look for physiological features that could predict whether a mental task will be correctly recognized by the BCI, and that could be understood by the user. Among the different features we explored, it appears that the user’s relaxation (from a muscular point of view), as measured in EMG activity collected by EEG channels, is one of such features. We are currently building and exploring new BCI training protocols that provide additional information about the user’s muscular relaxation as complementary feedback [34].

7.10. EEG Signal Processing

Participant: Fabien Lotte.

To make BCI practical and useful, we need to make them reliable, i.e., able to recognize the users’ mental commands, despite noise and non-stationarities [42]. We also need to reduce their calibration time, as current systems need many examples from each user to calibrate the system for this specific user. This year we addressed these two issues with two different studies.

In order to reduce BCI calibration times, we first surveyed existing approaches, these approaches being notably based on regularization, user-to-user transfer, semi-supervised learning and a-priori physiological information. We then proposed new tools to reduce BCI calibration time. In particular, we proposed to generate artificial EEG trials from the few EEG trials initially available, in order to augment the training set size. These artificial EEG trials are obtained by relevant combinations and distortions of the original trials available. We proposed 3 different methods to do so. We also proposed a new, fast and simple approach to perform user-to-user transfer for BCI. Finally, we studied and compared offline different approaches, both old and new ones, on the data of 50 users from 3 different BCI data sets. This enabled us to identify guidelines about how to reduce or suppress calibration time for BCI [16].

In order to increase BCI robustness, we performed an empirical comparison of covariance matrix averaging methods for EEG signal classification. Indeed, averaging EEG signal covariance matrices is a key step in designing brain-computer interfaces (BCI) based on the popular common spatial pattern (CSP) algorithm. BCI paradigms are typically structured into trials and we argue that this structure should be taken into account. Moreover, the non-Euclidean structure of covariance matrices should be taken into consideration as well. We reviewed several approaches from the literature for averaging covariance matrices in CSP and compared them empirically on three publicly available data sets. Our results showed that using Riemannian geometry for averaging covariance matrices improves performances for small dimensional problems, but also the limits of this approach when the dimensionality increases [36].

7.11. ECoG-based analysis of Speech processes

Participant: Fabien Lotte.

Acoustic speech output results from coordinated articulation of dozens of muscles, bones and cartilages of the vocal mechanism. While we commonly take the fluency and speed of our speech productions for granted, the neural mechanisms facilitating the requisite muscular control are not completely understood. Previous neuroimaging and electrophysiology studies of speech sensorimotor control has typically concentrated on speech sounds (i.e. phonemes, syllables and words) in isolation; sentence-length investigations have largely been used to inform coincident linguistic processing. In this study, we examined the neural representations of segmental features in the context of fluent, continuous speech production. We used recordings from the cortical surface (electrocorticography (ECoG)) to simultaneously evaluate the spatial topography and temporal dynamics of the neural correlates of speech articulation that may mediate the generation of hypothesized gestural or articulatory scores. We found some aspects of speech production (place of articulation) involved broad networks of brain regions during all phases of speech production: preparation, execution and monitoring. Other aspects (manner of articulation and voicing status) were dominated by auditory cortical responses after speech had been initiated. These results provide a new insight into the articulatory and auditory processes underlying speech production in terms of their motor requirements and acoustic correlates (see Figure 9 , [15]).

7.12. Toward a portable tangible EEG interface

Participants: Maxime Duluc, Thibault Laine, Jérémy Frey, Renaud Gervais, Fabien Lotte, Martin Hachet.

Last year we presented **Teegi**, the first interface that combines electroencephalographic (EEG) recordings and tangible interaction in order to let novices learn about how their brain works. By displaying EEG activity in real time on a support that is easy to manipulate and to comprehend, Teegi is a good tool for scientific outreach, that raises public interest.

Yet, the gap between research projects and the field is not often bridged. While our past prototype used an external projector and a supplementary tracking device to display information onto the head of the puppet, over the course of the year we developed a semi-spherical display based on LEDs (see Figure 10). By embedding all the electronics into the puppet, Teegi will be easier to bring outside the laboratory. Thanks to these technological advances, real-life applications of the system are finally within reach.

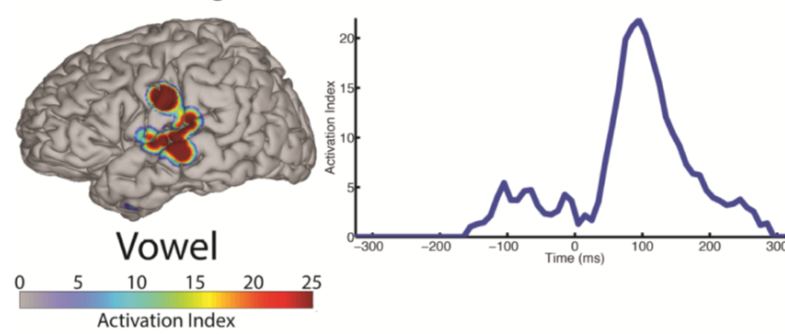


Figure 9. Exemple of the ECoG signature of vowel phonemes.

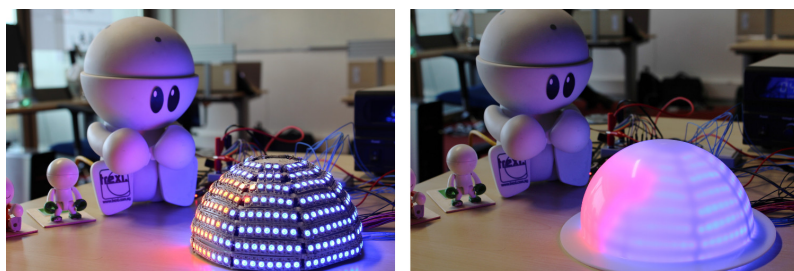


Figure 10. Development version of a portable version of Teegi, a Tangible EEG Interface. An array of LEDs and a diffuser replace the use of spatial augmented reality.

7.13. Electroencephalography-based evaluation of user experience

Participants: Jérémy Frey, Maxime Daniel, Dennis Wobrock, Julien Castet, Martin Hachet, Fabien Lotte.

Designing user interfaces requires adequate evaluation tools to ensure good usability and user experience. While many evaluation tools are already available and widely used, existing approaches generally cannot provide continuous and objective measures of usability qualities during interaction without interrupting the user. On the other hand, the measure of brain activity by the mean of electroencephalography (EEG) is mature enough to assess mental states. Combined with existing methods, such tool can be used to strengthen the understanding of user experience.

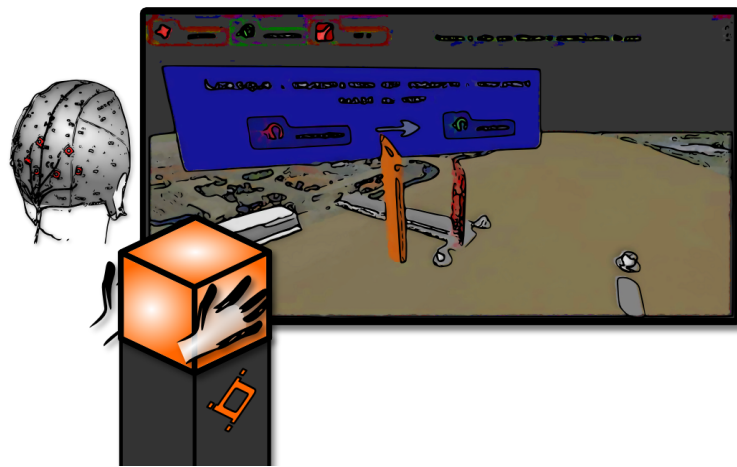


Figure 11. Schematic view of a user performing 3D manipulations tasks with the CubTile input device. His/her mental effort are monitored based on brain signals (electroencephalography).

In [35] we studied 3D object manipulation tasks. We showed how mental workload can be estimated from EEG, and then measured it on 8 participants during an actual 3D object manipulation task with an input device known as the CubTile (see figure 11). These first results suggested that we could continuously assess the 3DUI and/or interaction technique ease-of-use.

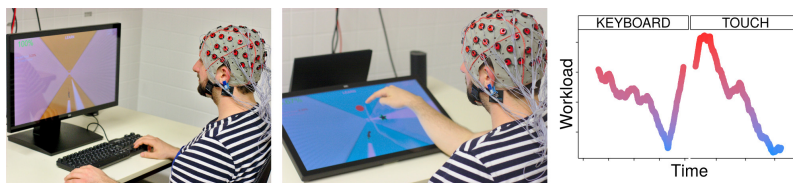


Figure 12. A keyboard (left) can be compared with a touch interface (middle) using a continuous measure of mental workload through electroencephalography (right).

We pushed further these finding in a second study [26], where we have developed a set of methods to continuously estimate the user's mental workload, attention level and recognition of interaction errors during

different interaction tasks. We validated these measures in a controlled virtual environment and showed how they can be used to compare different interaction techniques – for instance a keyboard and a touch-based interface (see Figure 12).

Thanks to such framework, EEG becomes a useful addition to the repertoire of available evaluation tools, enabling a finer grain assessment of the ergonomic qualities of computer systems.

7.14. Classifying EEG Signals during Stereoscopic Visualization to Estimate Visual Comfort

Participants: Jérémy Frey, Aurélien Appriou, Fabien Lotte, Martin Hachet.

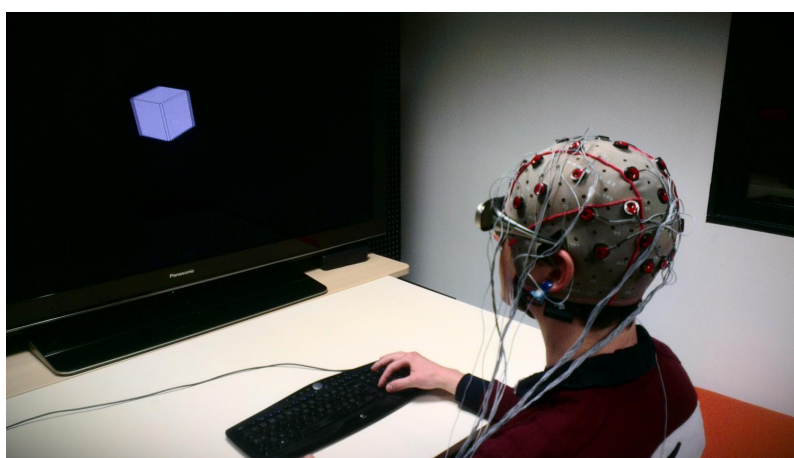


Figure 13. Setup of the experiment, with a subject being presented with stereoscopic images while his EEG signals are being recorded.

With stereoscopic displays a sensation of depth that is too strong can impede visual comfort and may result in fatigue or pain. We used Electroencephalography (EEG) to develop a novel brain-computer interface that monitors users' states in order to reduce visual strain. We present the first system that discriminates comfortable conditions from uncomfortable ones during stereoscopic vision of still images using EEG [13], [25] – see Figure 13). In particular, we show that changes in event-related potentials' (ERPs) amplitudes following stereoscopic objects presentation can be used to estimate visual comfort. Our system reacts within 1 second to depth variations, achieving 63% accuracy on average (up to 76%) and 74% on average when 7 consecutive variations are measured (up to 93%). Performances are stable ($\approx 62.5\%$) when a simplified signal processing is used to simulate online analyses or when the number of EEG channels is lessened. This study could lead to adaptive systems that automatically suit stereoscopic displays to users and viewing conditions. For example, it could be possible to match the stereoscopic effect with users' state by modifying the overlap of left and right images according to the classifier output

PRIMA Project-Team

6. New Results

6.1. Attention-Based Navigation

Participants: Thierry Fraichard, Remi Paulin, Patrick Reignier.

The domain of service-robots is growing fast and has become the focus of many researchers and industrials alike. Application areas are extremely broad, from logistics to handicap assistance. A large proportion of such robots are expected to share humans' living space and thus must be endowed with navigation capabilities that exceed the standard requirements pertaining to autonomous navigation such as motion safety. In a human populated environment, optimality does not boil down to minimizing resources such as time or distance traveled anymore, the robot motion must abide by social rules and move in a manner which is **appropriate**.

Most of the approaches proposed so far rely upon the definition of so-called **social spaces**, *i.e.* regions in the environment that, for different reasons, the persons consider as psychologically theirs. Such social spaces are primarily characterized using either the position of the person, *e.g.* "Personal space" [36], or the activity he is currently engaged in, *e.g.* "Interaction Space" [41] and "Activity Space" [45]. The most common approach is then to define costmaps on such social spaces: the higher the cost, the less desirable it is for the robot to be at the corresponding position. The costmaps are ultimately used for motion planning and navigation purposes.

While improving upon the standard "non social" navigation methods, this type of approach intrinsically ignores the correlations between interactions as well as the influence of the robot on those interactions. It thus fails to capture several important features of social navigation, such as the distraction and surprise caused to the surrounding individuals. To overcome those limits, we suggest using the psychological concept of attention, which plays a central role when humans navigate around each other. This concept brings a new degree of control over the motion of the robot, namely the invasive and distracting character of the robot motion, which have so far proven hard to tackle with the conventional tools such as social spaces. Beside leading appropriate motion, attention-based navigation enables interaction through motion by predicting the quantity of attention the human will give to the robot.

Building upon a computational model of attention that was earlier proposed in [47], we have developed the novel concept of **attention field**. The attention field is straightforward to define: it is a measure of the amount of attention that a given person would allocate to the robot, should the robot be in a given position/state. It is a mapping from the state space of the robot to \mathbb{R} . We use this attention field in order to carefully control the degree of distraction caused by the robot to the individuals in its surroundings. By monitoring the variations of attentional resources that it causes, we also control the amount of surprise caused by the robot which must be kept to a minimum since it is a cause of discomfort. Furthermore this approach enables us to tackle more complex situations where more than one person is involved such as the task of delivering a private message to an individual, or else joining a group (an example of interaction through motion). Rather than navigating on a single global costmap, this new approach provides for each path several measures of the distraction and surprise caused by the robot on a given individual. Those quantities are then multi-optimised in order to find a path that satisfies all the given requirements for fulfilling the robot's task as well as minimizing the discomfort for individuals who are not directly involved in an interaction with the robot.

In 2015, we have developed a variant of the well-known differential evolution algorithm which deals with optimizing continuous trajectories under multiple constraints. The performance of our approach is now being compared with trajectories obtained by relying only on social spaces. Besides the traditional qualitative approaches to evaluate the discomfort caused by the robot motion, we work on defining more quantitative measures that would enable us to further validate our approach.

6.2. SPOK: End User Programming for Smart Homes

Participant: Alexandre Demeure.

As part of the CATRENE project AppsGate, we have developed SPOK, an End User Development Environment, that enables inhabitants to control and program their smart Homes via a web interface. The current version of SPOK includes an editor for editing programs using a pseudo-natural language and an interpreter. A multi-syntax editor as well as additional services such as a debugger and a simulator are expected for the second version.

A multi-syntax editor will allow users to build syntactically correct programs using the syntax that is most appropriate to them or by using a combination of them. These syntaxes include pseudo-natural language (i.e. a constrained natural language) and graphical iconic syntax (as exemplified by Scratch [Maloney et al. 2010]). The interaction techniques used to enter programs may be menu-based, free typing, as well as by demonstration in the physical home or by the way of the simulator. The simulator is the dual digital representation of the real home. It is intended to serve also as a debugger for testing and correcting end-user programs.

Whatever syntax used by end-users, programs are translated into syntactic abstract trees whose leaves reference services provided by the Core HMI and/or by the Extended HMI Middleware. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input.

In order to support a dynamically extensible grammar as well as to provide end-users with feedforward at the user interface of the editor, the grammar used by the editor is split into 2 parts: the root grammar and the device specific grammars. The root grammar specifies the generic structures of an end-user program: loops, conditions, etc. The device specific grammars are separated from the root grammar to be able to dynamically build the final grammar to be compliant with what is currently installed and detected by the AppsGate server. Each device type brings with it its own events, status and actions. These grammatical elements are injected into the root grammar when generating the parser and for compiling end-user programs.

The language used by end-users to express their programs is a pseudo-natural language using the rule-based programming paradigm. The left hand side of a rule is composed of events and conditions, and the right hand side specifies the actions to be taken when the left hand-side is true or becomes true. A program may include several rules that can be executed either in parallel or sequentially. Once entered, programs are translated into syntactic abstract trees. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input. SPOK is implemented as a mix of OSGi and ApAM components where ApAM is in turn a middleware that runs on top of OSGi.

6.3. Qualitative approaches for building energy management

Participant: Patrick Reignier.

Reducing housing energy costs is a major challenge of the 21st century. In the near future, the main issue for building construction is the thermal insulation, but in the longer term, the issues are those of renewable energy (solar, wind, etc.) and smart buildings. Home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes. Thanks to this network, a load management mechanism can be carried out: it is called distributed control. An optimal home energy management system is still a goal to aim for, because lots of aspects are still not completely fulfilled. Most of the energy systems respect only the energy needs, but they don't tackle the user needs or satisfaction. Energy systems also have a lack when it comes to the dynamicity of the environments (the system ability to adapt). The problem is similar for the existing HMI (Human User Interface) of those Home Automation Systems where only experts can understand the data coming from the sensors and most important, the energy plan coming from management system (How? and Why?). The goal of this study is to propose a house energy model that can be both used to predict at some level energy evolution and that can be understood by the end user. The house energy model is based on Fuzzy Cognitive Maps representing cause-effects relations. It is first designed by an expert and then automatically tuned to a particular house using machine learning approaches. Preliminary experiments have been done this year using the Predis datasets.

6.4. Situation Aware Services on Mobile Devices

Participants: James Crowley, Thibaud Flury.

Modern mobile devices, such as smart phones and tablets, combine a rich set of sensors, internet connectivity, with embedded computational power and memory. The PRIMA group has recently demonstrated that it is possible to construct embedded software that uses the full suite of mobile sensors to recognise activities and learn the daily routines of users.

A first proof of concept has recently been constructed using recognition of places and activities. The system was trained by having student volunteers carry a cell phone running a data acquisition program that recorded signals from accelerometer, gyroscope, ambient sound, ambient light, Cell tower, wifi, bluetooth, and GPS based geolocalisation. The data were labeled by the students with ground truth data about transportation modes, places, and activities. This data was then used to learn recognition routines. Recognition of places, activities, and transportation was used to construct probabilistic models of daily routines using PRIMA's situation modelling techniques, previously demonstrated in constructing situation aware services. The system was demonstrated by constructing a Twitter Bot (a robot that publishes on twitter) that published information about volunteers during their daily activity.

A professional quality software system named CAM - Context Aware Manager - is currently under construction, and will be licensed to the PRIMA startup *Situ8ed*, for use in context aware mobile services.

6.5. Perceiving mass in mixed reality

Participants: Sabine Coquillart, Paul Issartel.

In mixed reality, real objects can be used to interact with virtual objects. However, unlike in the real world, real objects do not encounter any opposite reaction force when pushing against virtual objects. The lack of reaction force during manipulation prevents users from perceiving the mass of virtual objects. Although this could be addressed by equipping real objects with force-feedback devices, such a solution remains complex and impractical. In this work, we present a technique to produce an illusion of mass without any active force-feedback mechanism. This is achieved by simulating the effects of this reaction force in a purely visual way. A first study demonstrates that our technique indeed allows users to differentiate light virtual objects from heavy virtual objects. In addition, it shows that the illusion is immediately effective, with no prior training. In a second study, we measure the lowest mass difference (JND) that can be perceived with this technique. The effectiveness and ease of implementation of our solution provides an opportunity to enhance mixed reality interaction at no additional cost.

6.6. Pseudo-haptic feedback

Participants: Sabine Coquillart, Jingtao Chen.

"Pseudo-haptic feedback" is a technique aiming to simulate haptic sensations without active haptic feedback devices. Pseudo-haptic techniques have been used to simulate various haptic feedbacks such as stiffness, torques, and mass. In the framework of Jingtao Chen PhD thesis, a novel pseudo-haptic experiment has been set up. The aim of this experiment is to study the EMG signals during a pseudo-haptic task. A stiffness discrimination task similar to the one published in Lecuyer's PhD thesis has been chosen. The experimental set-up has been developed, as well as the software controlling the experiment. Pre-tests are under way. They will be followed by the tests with subjects.

RITS Project-Team

7. New Results

7.1. 2D Laser Based Road Obstacle Classification for Road Safety

Improvement

Participants: Pierre Merdrignac, Evangeline Pollard, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality to the vulnerable road users (VRU), indicating a strong need to protect such persons. Laser sensors have been extensively used for moving obstacles detection and tracking. Laser impacts are produced by reflection on these obstacles which suggests an information is available to recognize multiple road obstacles classes (pedestrian, cyclists, vehicles,...). We introduce a new system to address this problem that is divided in three parts: definition of geometric features describing road obstacles, multi-class object classification from an adaboost trained classifier and Bayesian estimation of the obstacle class. This approach benefits from consecutive observations of a single obstacle to estimate its class more precisely. We tested our system on some laser sequences and showed that it can estimate the class of some road obstacles around the vehicle with an accuracy of 87.4%. The vehicle class is determined with more than 97% of success. However, the main source of confusion is for static obstacles (posts and trees) for which 15% are classified as pedestrians. More detail can be found in [36], [16].

7.2. On line Mapping and Global Positioning technique based on evidential SLAM

Participants: Guillaume Trehard, Evangeline Pollard, Fawzi Nashashibi.

Locate a vehicle in an urban environment remains a challenge for the autonomous driving community. By fusing information from a LIDAR, a Global Navigation by Satellite System (GNSS) and the vehicle odometry, we introduced and developed an original solution based on evidential grids and a particle filter to map the static environment and simultaneously estimate the position in a global reference at a high rate and without any prior knowledge (see [39]).

7.3. PML-SLAM

Participants: Zayed Alsayed, Fawzi Nashashibi, Anne Verroust-Blondet.

Our goal is to improve localization systems performances in order to be able to navigate in large-scale urban environments. In this context, we first optimized CPU and memory consumption of a SLAM laser-based technique [52] by introducing a map manager system. This strategy allows a smooth navigation while saving and loading probabilities-grid submaps into/from a hard-disc in a transparent way (cf. [27]). This work was validated and extended in the context of ITS Bordeaux demonstrations (VEDECOM demonstrator), where GPS information was integrated into SLAM environment Maps.

7.4. Motion planning techniques

Participants: David Gonzalez Bautista, Fernando Garrido Carpio, Joshué Pérez Rastelli, Vicente Milanés Montero, Fawzi Nashashibi.

Intelligent vehicles have increased their capabilities for highly, and even fully, automated driving under controlled environments. Scene information is received using on-board sensors and communication network systems—i.e. infrastructure and other vehicles. Considering the available information, different motion planning techniques have been implemented to autonomously driving on complex environments. The main goal is focused on executing strategies to improve safety, comfort and energy optimization. However, research challenges such as navigation in urban dynamic environments with obstacle avoidance capabilities—i.e. Vulnerable Road Users (VRU) and vehicles—and cooperative maneuvers among automated and semi-automated vehicles still need further efforts for a real environment implementation. We have recently carried out a deep state-of-the-art review to find the gaps in this hot topic into the autonomous vehicle field, paying special attention to overtaking and obstacle avoidance maneuvers.

Based on this review, we have mainly identified two main gaps: trajectory and speed planning with dynamics obstacle avoidance capabilities and real-time performance of the algorithms in the sense of significantly reducing the computational time, moving the system closer to what a vehicle should be able to provide in the real world.

According to this review, a speed planner has been designed with specific considerations on computing time efficiency, with an optimal comfort and avoiding to exceed speed and acceleration limits [31]. The comfort is evaluated as the minimization and smoothness of acceleration and jerk profiles, while maintaining a coherent speed profile with respect to traffic rules, the geometry of the path and the lateral accelerations associated to it. Specifically, this speed planner uses fifth order polynomial curves. These curves are C2 continuous and smooth, meaning that the jerk profile is also continuous and smooth. The method proposed computes the velocity in terms of the length of the path, instead of time, greatly reducing the errors. Specific targets for the speed planner are:

- Compute a smooth and continuous speed profile accounting for acceleration limits (longitudinal and lateral) according to ISO 2631-1 standard.
- Minimize distance error problems by associating the speed profile in the path speed planner instead of the time.

This speed planner was tested against other techniques providing better results in terms of computational time and smoothness (cf. [32]).

Additionally, a novel trajectory planning with a significant reduction on the computational time with respect to prior implementations from the team has been implemented. Our approach is mainly affected by vehicle's kinematics and physical road constraints. Based on these assumptions, computational time for path planning can be significantly reduced by creating a database containing already optimized versions of all the potential trajectories in each curve the vehicle can carry out. Therefore, this algorithm generates a database of smooth and continuous curves considering a big set of different intersection scenarios, taking into account the constraints of the infrastructure and the physical limitations of the vehicle. According to the real scenario, the local planner selects from the database the appropriate curves, searching for the ones that fit with the intersections defined on it. The path planning algorithm has been tested in simulation against the previous control architecture. The results obtained show path generation improvements in terms of smoothness and to continuity. Next steps on this algorithm is to test its performance in real platform and add the dynamics obstacle avoidance capabilities, establishing the link with the perception algorithms research line currently open in the team.

7.5. Control techniques

Participants: Francisco Navas Matos, Carlos Eduardo Flores Pino, David Gonzalez Bautista, Joshué Pérez Rastelli, Vicente Milanés Montero.

The final stage for automating a vehicle relies on the control algorithms. They are in charge of providing the proper behavior and performance to the vehicle, leading to provide the fully automated capabilities. Having this in mind, there are two research lines currently open in the time: the first one is mainly related to what we call “naturalistic driving” in the sense of adding the human reasoning to the vehicle. We are mainly focusing

our effort on artificial intelligent algorithms as neuro-fuzzy techniques. The main reason is the growing interest of the car makers in adding sharing control capabilities (between the vehicle and the driver) to the automated car. Our initial results show a big potential of using this approach and we already achieved some simulations results that were well-accepted by the scientific community and will be shown in mid-December at the final event of the EU project DESERVE.

On the other hand, we are also further investigating robust control algorithms for providing stability not only to an automated vehicle but also to a chain of automated vehicles that should be able to cooperate intelligently. This work is mainly divided in two main research lines:

1) Controllability and stability of dynamic complex systems are the key aspects when it comes to design intelligent control algorithms for vehicles. Current advances in the field are mainly oriented to advanced multi-sensor fusion toward multi-target decision-making systems. These artificial intelligence-based algorithms are able to provide reasonable responses under controlled environments (i.e. highly-detailed maps). However, new trends are proposing intelligent algorithms able to handle any unexpected circumstances as unpredicted uncertainties or even fully outages from sensors. The goal of this new research line at RITS is to further investigate control algorithms able to provide stability responses for autonomous vehicles under uncontrolled circumstances, including modifications on the input/output sensors. Dynamic plant models where different inputs/outputs can be added or subtracted in real-time during its operation is one of the hot topics in the control research arena. This system has to provide stable enough response when these operations occur. This is especially true on high-risk environments as autonomous driving; and

2) Data-driven control techniques based on model-free algorithms. Vehicles exhibit a highly non-linear behavior, especially at low speeds (as occur in urban environments). The research on novel data-driven techniques that are independent of the plant model provides huge benefits when applying them to automated vehicles. This novel research line in the team tries to further investigate on stable algorithm that doesn't need an accurate model of the vehicle dynamic, leading to compensate the effects of nonlinear dynamics, disturbances, or uncertainties in the parameters. [35]

7.6. Study on Perception and Communication Systems for Safety

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Ines Ben Jemaa, Fawzi Nashashibi.

The existing R&D efforts for protecting vulnerable road users (VRU) are mainly based on perception techniques, which aim to detect VRUs utilizing vehicle embedded sensors. The efficiency of such a technique is largely affected by the sensor's visibility condition. Vehicle-to-Pedestrian (V2P) communication can also contribute to the VRU safety by allowing vehicles and pedestrians to exchange information. This solution is, however, largely affected by the reliability of the exchanged information, which most generally is the GPS data. Since perception and communication have complementary features, we can expect that a combination of such approaches can be a solution to the VRU safety. This is the motivation of this work. We develop theoretical models to present the characteristics of perception and communications systems. Experimental studies are conducted to compare the performances of these techniques in real-world environments. Our results show that the perception system reliably detects pedestrians and other objects within 50 m of range in the line-of-sight (LOS) condition. In contrast, the V2P communication coverage is approximately 340 and 200 meters in LOS and non-LOS (NLOS) conditions, respectively. However, the communication-based system fails to correctly position the VRU w.r.t the vehicle, preventing the system from meeting the safety requirement. Finally, we propose a cooperative system that combines the outputs of the communication and perception systems. More detail can be found in [37], [16].

7.7. Asynchronous Reactive Distributed Congestion Control Algorithms for the ITS G5 Vehicular Communications

Participant: Oyunchimeg Shagdar.

The IEEE 802.11p is the technology dedicated to vehicular communications to support road safety, efficiency, and comfort applications. A large number of research activities have been carried out to study the characteristics of the IEEE 802.11p. The key weakness of the IEEE 802.11p is the channel congestion issue, where the wireless channel gets saturated when the road density increases. The European Telecommunications Standardization Institute (ETSI) is in the progress of studying the channel congestion problem and proposed so-called Reactive Distributed Congestion Control (DCC) algorithm as a solution to the congestion issue. In this work we investigate the impacts of the Reactive DCC mechanism in comparison to the conventional IEEE 802.11p with no congestion control. Our study shows that the Reactive DCC scheme creates oscillation on channel load that consequently degrades communication performance. The results reveal that the channel load oscillation is due to the fact that in the Reactive DCC, the individual CAM (Cooperative Awareness Message) controllers react to the channel congestion in a synchronized manner. To reduce the oscillation, we propose a simple extension to Reactive DCC, Asynchronous Reactive DCC, in which the individual CAM controllers adopt randomized rate setting, which can significantly reduce the oscillation and improve the network performance. See [45] for more detail.

7.8. Vehicle to vehicle visible light communication

Participants: Mohammad Abu Alhoul, Oyunchimeg Shagdar, Fawzi Nashashibi.

Visible Light Communication (VLC) technology utilizes the light spectral range between 380 nm and 750 nm, which enables the dual functionality of lightning and information delivery. A use of VLC for the ITS domain has many benefits including that it can be a complementary technology to the IEEE 802.11p, which is the radio communications technology dedicated to the V2X communication but suffers from its channel congestion problem.

This year, we conducted theoretical and experimental studies on the optical channel characteristics. Based on our studies and the previous contributions, we developed a transmitter and receiver VLC prototype to be integrated to the vehicle lightning systems dedicated to platooning applications. Using the low-cost Arduino micro-controller, a transmitter broadcasts the vehicle status information including the vehicle identity, velocity, orientation, acceleration through the vehicle rear Light Emitting Diodes (LED). The receiver is based on a simple Photo Diode (PD) with an accurate 635 nm optical filtering stage to overcome the saturation and the unwanted ambient noise issues. Experimental studies show that the system can provide 8.5 Kbps of information delivery between vehicles with up to 30 meters of bumper to bumper distance.

7.9. Analysis of broadcast strategies in IEEE 802.11p VANETs

Participants: Younes Bouchaala, Oyunchimeg Shagdar, Paul Muhlethaler.

We analyze different broadcast strategies in IEEE 802.11p Vehicular Ad-hoc NETWORKS (VANETs). The first strategy is the default IEEE 802.11p strategy. Using a model derived from the Bianchi model, we provide the network performance in terms of throughput and success rate. The second strategy consists in using an acknowledgment technique similar to the acknowledgment with point-to-point traffic. A node will send its broadcast packet as in the default case, but it requires an acknowledgment from a neighbor node. This node may be a random neighbor or may be selected according to precise rules. We analyze this second strategy in terms of throughput and success rate. Somewhat surprisingly, we show that this second strategy improves the delivery ratio of the transmitted packets but reduces the overall throughput. This means that if the CAM messages (Cooperative Awareness Messages) are broadcasted, the total number of packets actually delivered will be greater with the default strategy than with the improved strategy. We propose a third strategy which consists in using the default strategy for normal packets, but we add random redundant transmissions to ensure greater reliability for very important packets. We show that with this simple technique, not only do we obtain suitable reliability, but we also achieve larger global throughput than with the acknowledgment-oriented technique. This is described in [26]. Another contribution of this paper is to compute network performance in terms of throughput and success rate with respect to the network parameters and to analyze their impact on performances.

7.10. Multicast communications for cooperative vehicular systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Paul Muhlethaler, Arnaud de La Fortelle.

With the advancement of wireless communications technologies, users can now have multicast services while they are driving. Majority of the multicast services require Internet-to-vehicle multicast message dissemination. Conventional group management approaches in Internet is relatively simple because it is performed on the local networks of the multicast members which are usually a priori configured to receive the service. In addition to this, multicast packets flows follow a fixed routing structure that is built between the source and the destinations. These approaches could not be applied to vehicular networks (VANET) due to their dynamic and distributed nature. In order to enable such multicasting, our work deals with two aspects. First, reachability of the moving vehicles to the multicast service and second, multicast message dissemination in the VANET. Regarding the first issue, we find that neither current multicast addressing nor existing mobility management mechanisms are suitable for VANET. We introduce first a self-configuring multicast addressing scheme that allows the vehicles to auto-configure a dynamic multicast address without a need to exchange signalling messages with the Internet. Second, we propose a simplified approach that extends Mobile IP and Proxy Mobile IP. About message dissemination, we first propose to revisit traditional multicast routing techniques that rely on a tree structure. In particular, as vehicular networks are known to have changing topology, we present a theoretical study of the link lifetime between vehicles in urban environments. Then, we propose then Motion-MAODV, an improved version of a tree-based routing mechanism (MAODV) that aims at guaranteeing longer route lifetime. Finally, we also propose a geographic routing protocol Melody that provides a geocast dissemination in urban environments. Through simulations, we show that Melody ensures more reliable and efficient packet delivery to a given geographic area compared to traditional geo-broadcasting schemes in highly dense scenarios. More detail can be found in [28], [41], [47].

7.11. Context Awareness and Priority Control for ITS based on Automatic Speech Recognition

Participants: Oyunchimeg Shagdar, Sakriani Watiasri Sakti.

Bringing rapid assistance to motorists involved in a traffic accident is an important service to be provided by Intelligent Transportation System (ITS). Existing proposals to automatic accident detection are based on the vehicle's perception point of view. In [38] we introduce situational awareness based on the "understanding" of conversational speech of drivers/passengers using an automatic speech recognition (ASR) system. Context-aware priority control and congestion control schemes are presented to ensure coexistence of ASR-triggered applications and cooperative awareness messages (CAM) in the IEEE 802.11p system. Finally, application risk analysis and performance evaluations of ASR and V2X communications are carried out.

7.12. Emergent Behaviors and Traffic Density among Heuristically-Driven Intelligent Vehicles using V2V Communication

Participants: Oyunchimeg Shagdar, Fawzi Nashashibi.

We study the global traffic density and emergent traffic behavior of several hundreds of intelligent vehicles, as a function of V2V communication (for the ego vehicle to perceive traffic) and path-finding heuristics (for the ego vehicle to reach its destination), in urban environments. Ideal/realistic/no V2V communication modes are crossed with straight-line/towards-most-crowded/towards-least-crowded pathfinding heuristics to measure the average trip speed of each vehicle. The behaviors of intelligent vehicles are modeled by a finite state automaton. The V2V communication model is also built based on signal propagation models in an intersection scenario and a Markov-chain based MAC model. Our experiments in simulation over up to 400 vehicles exhibit attractive insights: 1) communication's impact is positive for the performance of the emergent vehicles' behavior, however, 2) the path-finding heuristics may not obtain their expected collective behavior due to the communications errors in realistic road environment (cf. [43]).

7.13. Time-bounded message dissemination in strings

Participant: Gérard Le Lann.

In 2015, besides reviewing prominent open issues regarding safety in IVNs (see [42]), we have investigated coordination problems that arise in string formations. Since the inception of the platoon concept (1977), a number of solutions have been proposed for achieving string control (platoons are a particular case of ad hoc/open string). String control must be exercised in order to avoid rear-end collisions, string instability, and for coping with emergency situations. The cyber components essential for string control have not been fully identified yet. For example, considering the cooperative adaptive cruise control paradigm, data collected in recent platooning experiments show that it is inappropriate to rely on V2V broadcast from a lead vehicle, thus the quest for other approaches. In strings, one can take advantage of short-range directional antennas which enable fast messaging among consecutive string neighbors, leading to the concept of neighbor-to-neighbor (N2N) communications and the cohort construct (a cohort is a string with a specification). String control problems translate into communication protocol issues and distributed algorithmic problems, notably:

- Time-bounded string-wide acknowledged message delivery and dissemination (TBMD),
- Bounded channel access delay (BCAD), a MAC-level problem,
- Time-bounded message acknowledgment (TBMA).

Acceptable solutions shall achieve small non-stochastic worst-case channel access time bounds (BCAD) and bounded delays for successful message delivery (TBMA and TBMD), under worst-case conditions regarding channel contention and message/acknowledgment losses. Non-stochastic worst-case bounds can only be established analytically (obviously, simulations cannot be considered). The importance of the TBMD problem can be exposed simply as follows: would TBMD be solved, then the string instability problem vanishes. Rather than resting solely on stepwise detection-and-reaction strategies based on radars/lasers, every string member adjusts its acceleration/deceleration rate according to observed motions of its predecessor, TBMD delivers a N2N message carrying the newly string-wide targeted velocity, in less than 100 milliseconds in strings comprising in the order of 20 members, in the presence of message/acknowledgment losses. The TBMD problem has been solved (see [34]). The solution rests on assuming that TBMA and TBMD have solutions. Both problems have been solved (solutions are under review). Contrary to strings, groups are ad hoc/open multilane formations. It turns out that solutions aimed at the 3 problems referenced above are instrumental in solving problems arising with multilane SC scenarios. For example, the 3-way handshakes at the core of safe lane changes published previously now achieve significantly better performance figures. Work in progress also includes:

- conflicting concurrent lane changes at high velocities,
- fully automated zipper merging at high velocities, in non-line-of-sight conditions (radio communications), in line-of-sight conditions (optical communications).

7.14. Broadcast Transmission Networks with Buffering

Participants: Guy Fayolle, Paul Muhlethaler.

We analyzed the so-called back-off technique of the IEEE 802.11 protocol in broadcast mode with waiting queues. In contrast to existing models, packets arriving when a station (or node) is in back-off state are not discarded, but are stored in a buffer of infinite capacity. As in previous studies, the key point of our analysis hinges on the assumption that the time on the channel is viewed as a random succession of transmission slots (whose duration corresponds to the length of a packet) and mini-slots during which the back-off of the station is decremented. These events occur independently, with given probabilities. The state of a node is represented by a two-dimensional Markov chain in discrete-time, formed by the back-off counter and the number of packets at the station. Two models are proposed both of which are shown to cope reasonably well with the physical principles of the protocol. Stability (ergodicity) conditions are obtained and interpreted in terms of maximum throughput. Several approximations related to these models are also discussed in [44].

7.15. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

This work [51] deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

These studies have been done in particular in the framework of the projects Travesti and Pumas.

This year, the work about the theoretical aspects of encoding real valued variables into a binary Ising model has been accepted for publication in *Annals of Mathematics and Artificial Intelligence* [23]. Moreover, an informal collaboration has been started with the company SISTeMA ITS, in order to assess the performance of our techniques in real-world city networks.

7.16. Random Walks in Orthants

Participant: Guy Fayolle.

7.16.1. Explicit criterion for the finiteness of the group in the quarter plane

In the book [3], original methods were proposed to determine the invariant measure of random walks in the quarter plane with small jumps, the general solution being obtained via reduction to boundary value problems. Among other things, an important quantity, the so-called *group of the walk*, allows to deduce theoretical features about the nature of the solutions. In particular, when the *order* of the group is finite, necessary and sufficient conditions have been given in [3] for the solution to be rational or algebraic. When the underlying algebraic curve is of genus 1, we propose, in collaboration with R. Iasnogorodski (St-Petersburg, Russia), a concrete criterion ensuring the finiteness of the group. It turns out that this criterion is always tantamount to the cancellation of a single constant, which can be expressed as the determinant of a matrix of order 3 or 4, and depends in a polynomial way on the coefficients of the walk [20].

7.16.2. Second Edition of the Book *Random walks in the Quarter Plane*

In collaboration with R. Iasnogorodski (St-Petersburg, Russia) and V. Malyshev, we prepared the second edition of the book [3], which will be published by Springer, in the collection *Probability Theory and Stochastic Processes*. Part II of this second edition borrows specific case-studies from queueing theory, and enumerative combinatorics. Five chapters will be added, including examples and applications of the general theory to enumerative combinatorics. Among them:

- Explicit criteria for the finiteness of the group, both in the genus 0 and genus 1 cases.
- Chapter *Coupled-Queues* shows the first example of a queueing system analyzed by reduction to a BVP in the complex plane.
- Chapter *Joining the shorter-queue* analyzes a famous model, where maximal homogeneity conditions do not hold, hence leading to a system of functional equations.
- Chapter *Counting Lattice Walks* concerns the so-called *enumerative combinatorics*. When counting random walks with small steps, the nature (rational, algebraic or holonomic) of the generating functions can be found and a precise classification is given for the basic (up to symmetries) 79 possible walks.

7.17. Global optimization for online resource allocation

Participant: Jean-Marc Lasgouttes.

As part of the Mobility 2.0 FP7 project, we have considered the possibility to allocate charging stations to Full Electric Vehicle (FEV) users in a way that, instead of merely minimizing their travel time, tries to improve the travel time for the whole community.

Our setting can be seen as a resource allocation problem, known as the Transportation Problem in Operations Research literature. It is solvable using several algorithms, among which the simplex algorithm or the Hungarian algorithm. Unfortunately, these algorithms are not well-adapted here for two reasons:

- The allocation of slots to users is done on-line, when the user does a request. It is not possible to wait until all the users are known before doing the allocation;
- The complexity of these algorithms is very high, especially since, due to the effect of range limitations, each request has different characteristics, which is equivalent to increasing the types of customers.

We therefore present a simple heuristic approach, which is fast enough for systems with thousands of stations. Its principle is to penalize the cost for the user with an approximation of the extra cost incurred to future users who compete for the same resource (a charging or parking slot).

This work has been presented at the ITSC'2015 conference [33].

SEMAGRAMME Project-Team

6. New Results

6.1. Syntax-semantics interface

6.1.1. Lexical Semantics

The interpretation of natural language utterances relies on two complementary elements of natural language modeling. On the one hand, the description of the combinatorics of natural language expresses how elementary units, or *lexical units* (typically the word), combine in order to build more complex elements, such as sentences or discourses. On the other hand, the description of these elementary units specifies how they contribute to the meaning of the whole by their *lexical meaning*. This specification should also take into account how the different parts of the lexical meanings combine during the *composition* process and how they relate to their underlying meaning concepts. For instance, the verbs *buy* and *sell* should refer to a common conceptual representation. However, their syntactic arguments (e.g., the subject) play a different (semantic) role with respect to the *transaction* concept that they share.

The modeling of these concepts and how they relate to each other gave rise to Frames Semantics as a representation format of conceptual and lexical knowledge [40], [31], [26], [59]. Frames consists of directed graphs where nodes correspond to entities (individuals, events, ...) and edges correspond to (functional or non-functional) relations between these entities. Providing a fine-grained representation of the internal concept structure allows both for a *decomposition* of the lexical meaning and for a precise description of the sub-structural interactions in the semantic composition process [58].

Frames can be formalized as extended typed feature structures [71], [50] and specified as models of a suitable logical language. Such a language allows for the composition of lexical frames on the sentential level by means of an explicit syntax-semantics interface [50]. Yet, this logical framework does not provide a direct link between Frames and truth-conditional semantics, where natural language utterances are considered with respect to the conditions under which they are true or false. In particular, it does not provide means for the lexical items to introduce explicit quantification over entities or events.

To overcome these limitations, we proposed use Hybrid Logic (HL) [27], [25]. HL is an extension of modal logic. As such, it is well-suited to the description of graph structures. Moreover, HL introduce *nominals*, that allow the logical formulas to refer to specific nodes of the graph. It is then possible, for example, to specify when two edges should meet. Moreover, it introduces *variables* for nodes, and the associated *quantifiers*, that can appear in the logical formulas. We used this framework to model quantification in Frame Semantics [23], [18]

6.1.2. Compositionality and Modularity

One says that a semantics is compositional when it allows the meaning of a complex expression to be computed from the meaning of its constituents. One also says that a system is modular if it is made of relatively independent components. In the case of a semantic system (e.g. a Montague grammar), we say that it is modular if the ontology on which it is based (including notions such as *truth*, *entities*, *events*, *possible worlds*, *time intervals*, *state of knowledge*, *state of believe*, ...) is obtained by combining relatively independent simple ontologies.

The intensionalization procedure introduced in [4] provides a first step towards modularity. It allows the extensional interpretation of a language to be transformed into an intensionalized interpretation that offers room for accommodating truly intensional phenomena. Moreover, this procedure is conservative in the sense that it preserves the truth conditions of sentences. Another instance of such a procedure is provided by the dynamization procedure described in [57], which allows a static interpretation to be turned into a dynamic one capable of accommodating phenomena related to discourse dynamics.

In [15], we showed that both the intensionalization and dynamization procedures are instances of an abstract general scheme for which conservativity results may be established using the notion of logical relation.

6.1.3. Abstract Categorical Parsing

Kanazawa [53], [54] has shown how parsing and generation may be reduced to datalog queries for a class of grammars that encompasses mildly context-sensitive formalisms. These grammars, which he calls *context-free λ -term grammars*, correspond to second-order abstract categorical grammars.

In [14], we showed how Kanazawa's reduction may be carried out in the case of abstract categorical grammars of a degree higher than two. To this end, we reduced the parsing problem for general Abstract Categorical Grammars to a provability problem in Multiplicative Exponential Linear Logic.

6.2. Discourse dynamics

6.2.1. Discourse Structure Modeling

It is usually assumed that the internal structure of a text, typically characterized by discourse or rhetorical relations, plays an important role in its overall interpretation. In order to build such a structure, some approaches rely on discourse grammars. The key idea is to consider the structural regularities in discourse structure similarly as syntactic regularities. A particular trend relies on tree grammars. This trend has been further developed by integrating the modeling of both clausal syntax and semantics, and discourse syntax and semantics within the frameworks of Tree-Adjoining Grammar (TAG) [48], [49] and TAG for Discourse (D-LTAG) [79], [41], [80], [42].

Two important features characterize these approaches. First, while they use a single grammatical formalism, two different grammars are used for syntactic parsing and then for discourse parsing. In addition to adding an intermediate processing step, this two-tiered treatment both complicates the modeling of connectives that are ambiguous in their syntactic and discourse use, and prevents using standard disambiguation techniques. Second, some discourse structures better represented by directed acyclic graphs (DAG) than by trees are not accounted for.

In order to address the second issue of building DAG structures, [36], [37] have proposed Discourse Synchronous TAG (D-STAG), a TAG based approach together with a higher-order interpretation of sentences using Synchronous Tree-Adjoining Grammar (STAG) [67], [77].

We developed a method to interface a sentential grammar and a discourse grammar without resorting to an intermediate processing step. The method is general enough to build discourse structures that are DAG and not only trees. Our analysis is based on D-STAG. We also use an encoding of TAG into ACG. This encoding allows us to express a higher-order semantic interpretation that enables building DAG discourse structures on the one hand, and to smoothly integrate the sentential and the discourse grammar thanks to the modular capability of ACG. The results has been published [13] and all the examples of the article have been implemented and may be run and tested with the ACGtk software (see 5.1).

6.2.2. Effects and Handlers

We made the argument that pragmatics are to semantics what side effects are to calculations in a programming language. We demonstrated this parallel on two aspects.

First off, both pragmatics and side effects serve the same function. Side effects in programming languages account for the effects of expressions that reach beyond their scope and for the way a language interacts with the world of its users. Pragmatics is concerned with phenomena that also involve the non-immediate effects of expressions (e.g., discourse anaphora, presupposition accommodation) and with the way language interacts with the world of its users. Secondly, we pointed out that very similar formal theories are being used to treat the both of them (i.e. monads and continuations).

Having established this parallel, we then put forward a preliminary proposal of integrating semantics and pragmatics while keeping them separate by assigning effectful computations of truth values as meanings of linguistic expressions. In this way, we can implement the pragmatics at the level of the side effects and then focus on pure semantics at the level of values.

6.3. Common basic resources

6.3.1. Graph Rewriting

Bruno Guillaume and Guy Perrier have proposed to use Graph Rewriting for parsing syntactic dependencies [17]. It is an application of a Graph Rewriting formalism that they have established with Guillaume Bonfante and Mathieu Morey [32] and implemented in the Grew software [47]. They have developed a system of rewriting rules dedicated to French, which they have evaluated by parsing the Sequoia corpus [33].

6.3.2. Categorical Logic

Elaborating on the work of Grishin [45], Moortgat has introduced the non-associative Lambek-Grishin calculus (LG) as the foundations of a new kind of symmetric categorial grammar [63], [64], which allows for the treatment of linguistic phenomena such as displacement or discontinuous dependencies.

In [16], we compared LG with the non-associative classical Lambek calculus (CNL) introduced by de Groote and Lamarche [81]. We provided a translation of LG into CNL, which allows CNL to be seen as a non-conservative extension of LG. We then introduced a bimodal version of CNL that we called 2-CNL. This allowed us to define a faithful translation of LG into 2-CNL. Finally, we showed how to accommodate Grishin's interaction principles by using an appropriate notion of polarity. From this, we derived a new one-sided sequent calculus for LG.

6.3.3. Deep Syntax Annotation of the Sequoia French Treebank

Deep-sequoia introduces a deep syntactic representation scheme for French, built from the surface annotation scheme of the Sequoia corpus and abstracting away from it [69]. This scheme expresses the grammatical relations between content words. When these grammatical relations take part into verbal diatheses, the diatheses are considered as resulting from redistributions from the canonical diathesis, which is retained in the annotation scheme. The first version of the deep-sequoia corpus was released in 2014.

In November 2015, a new version (7.0) of the corpus was released (see <http://deep-sequoia.inria.fr>). Most of the modifications were corrections of annotations that improve the overall consistency of the corpus. Marie Candito and Guy Perrier have published the annotation guidelines associated with the corpus in [22].

6.3.4. Large Scale Grammatical Resources

Guy Perrier and Bruno Guillaume have achieved the development of a French grammar FRIGRAM with a large coverage [12] in the formalism of Interaction Grammars [5]. The originality of the formalism lies in its system of polarities which expresses the resource sensitivity of natural languages and which is used to guide syntactic composition. We present the principles underlying grammar design, highlight its modular architecture and show that the lexicon used is independent of the grammar formalism. We also introduce the "companion property", and show that it helps to enforce grammar consistency.

6.3.5. Universal Dependency Treebank

Bruno Guillaume participates with Marie-Catherine de Marneffe to the production of the French sub-corpus of the Universal Dependency Treebank [68]. In November 2015, the version 1.2 was released. On the French sub-corpus, Grew was used to detect inconsistency and to correct automatically systematic errors.

SIROCCO Project-Team

7. New Results

7.1. Analysis and modeling for compact representation and navigation

3D modelling, multi-view plus depth videos, Layered depth images (LDI), 2D and 3D meshes, epitomes, image-based rendering, inpainting, view synthesis

7.1.1. Visual attention

Participants: Pierre Buysens, Olivier Le Meur.

Visual attention is the mechanism allowing to focus our visual processing resources on behaviorally relevant visual information. Two kinds of visual attention exist: one involves eye movements (overt orienting) whereas the other occurs without eye movements (covert orienting). Our research activities deals with the understanding and modeling of overt attention as well as saliency-based image editing. These research activities are described in the following sections.

Saccadic model: Most of the computation models of visual attention output a 2D static saliency map. This single topographic saliency map which encodes the ability of an area to attract our gaze is commonly computed from a set of bottom-up visual features. Although the saliency map representation is a convenient way to indicate where we look within a scene, these models do not completely account for the complexities of our visual system. One obvious limitation concerns the fact that these models do not make any assumption about eye movements and viewing biases. For instance, they implicitly make the hypothesis that eyes are equally likely to move in any direction.

There is evidence for the existence of systematic viewing tendencies. Such biases could be combined with computational models of visual attention in order to better predict where we look. Such a model, predicting the visual scanpath of observer, is termed as saccadic model. We recently propose a saccadic model ([20]) that combines bottom-up saliency maps, viewing tendencies and short-term memory. The viewing tendencies are related to the fact that most saccades are small (less than 3 degrees of visual angle) and oriented in the horizontal direction. Figure 1 (a) illustrates the joint probability distribution of saccade amplitudes and orientations. Examples of predicted scanpaths are shown in Figure 1 (b). We demonstrated that the proposed model outperforms the best state-of-the-art saliency models.

In the future, the goal is to go further by considering that the joint distribution of saccade amplitudes and orientations is spatially variant and depends on the scene category.

Perceptual-based image editing: Since the beginning of October, we have started new studies related to perceptual-based image editing. The goal is to combine the modelling of visual attention with image/video editing methods. More specifically it aims at altering images/video sequences in order to attract viewers attention over specific areas of the visual scene. We intend to design new computational editing methods for emphasizing and optimizing the importance of pre-defined areas of the input image/video sequence. There exist very few studies in the literature dealing with this problem. Current methods simply alter the content by using blurring operation or by recoloring the image locally so that the focus of attention falls within the pre-defined areas of interest. One avenue for improving current methods is to minimize a distance computed between a user's defined visual scanpath and predicted visual scanpath. The content would be edited (i.e. recoloring, region rescaling, local contrast/resolution adjustment, removing disturbing object, etc) in an iterative manner in order to move the focus of attention towards the regions selected by the user.

7.1.2. Epitome-based video representation

Participants: Martin Alain, Christine Guillemot.

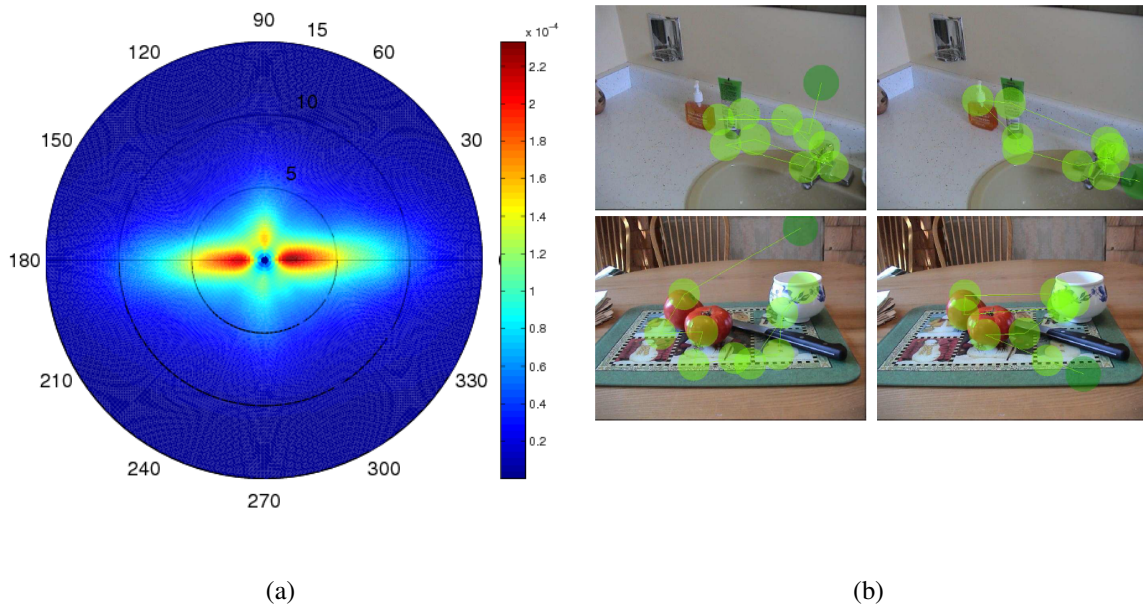


Figure 1. (a) Joint probability distribution of saccade amplitudes and orientations shown on a polar plot. Radial position indicates saccadic amplitudes expressed in degree of visual angle. (b) Predicted scanpaths composed of ten fixations represented by green circles. The dark green circle corresponds to the first fixation which is randomly chosen.

In 2014, we have developed fast methods for constructing epitomes from images. An epitome is a factorized texture representation of the input image, and its construction exploits self-similarities within the image. Known construction methods are memory and time consuming. The proposed methods, using dedicated list construction on one hand and clustering techniques on the other hand, aim at reducing the complexity of the search for self-similarities.

In 2015, we have developed methods for quantization noise removal (after decoding) exploiting the epitome representations together with local learning of either LLE (locally linear embedding) weights, which has proved to be a powerful tool for prediction [14], or using linear mapping functions between original and noisy patches. Compared to classical denoising methods which, most of the time, assume additive white Gaussian noise, the quantization turns out to be correlated to the signal which makes the problem more difficult. The methods have been experimented both in the contexts of single layer encoding and scalable encoding. The same methodology has been applied to super-resolution learning this time mapping functions between the low resolution and high resolution spaces in which lie the patches of the epitome [32].

7.1.3. Graph-based multi-view video representation

Participants: Christine Guillemot, Thomas Maugey, Mira Rizkallah, Xin Su.

One of the main open questions in multiview data processing is the design of representation methods for multiview data, where the challenge is to describe the scene content in a compact form that is robust to lossy data compression. Many approaches have been studied in the literature, such as the multiview and multiview plus depth formats, point clouds or mesh-based techniques. All these representations contain two types of data: i) the color or luminance information, which is classically described by 2D images; ii) the geometry information that describes the scene 3D characteristics, represented by 3D coordinates, depth maps or disparity vectors. Effective representation, coding and processing of multiview data partly rely on a proper representation of the geometry information. The multiview plus depth (MVD) format has become very popular in recent years for 3D data representation. However, this format induces very large volumes of data, hence the

need for efficient compression schemes. On the other hand, lossy compression of depth information in general leads to annoying rendering artefacts especially along the contours of objects in the scene.

Instead of lossy compression of depth maps, we consider the lossless transmission of a geometry representation that captures only the information needed for the required view reconstructions. Our goal is to transmit “just enough” geometry information for accurate representation of a given set of views, and hence better control the effect of geometry lossy compression.

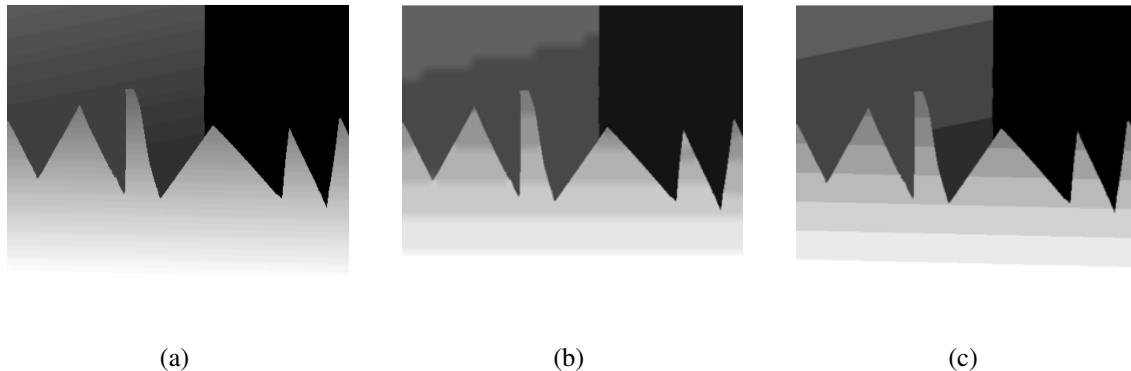


Figure 2. (a) original depth map, (b) depth map compressed with edge-adaptive method at 10kb with compression artifacts (c) depth image retrieved from the graph of our proposed GBR transmitted at 10kb keeping the original scene structure.

More particularly, in [23], we proposed a new Graph-Based Representation (GBR) for geometry information, where the geometry of the scene is represented as connections between corresponding pixels in different views. In this representation, two connected pixels are neighboring points in the 3D scene. The graph connections are derived from dense disparity maps and provide just enough geometry information to predict pixels in all the views that have to be synthesized.

GBR drastically simplifies the geometry information to the bare minimum required for view prediction. This “task-aware” geometry simplification allows us to control the view prediction accuracy before coding compared to baseline depth compression methods (Fig. 2). This work has first been carried out for multi-view configurations, in which cameras are parallel. We are currently investigating the extension of this promising GBR to complex camera transitions. An algorithm has already been implemented for two views and is being extended for multiple views. The next steps will be to develop color coding tools adapted to these graph structures.

7.2. Rendering, inpainting and super-resolution

image-based rendering, inpainting, view synthesis, super-resolution

7.2.1. Color and light transfer

Participants: Hristina Hristova, Olivier Le Meur.

Color transfer aims at modifying the look of an original image considering the illumination and the color palette of a reference image. It can be employed for image and video enhancement by simulating the appearance of a given image or a video sequence. It can also be applied to hallucinations of particular parts of the day. Current state-of-the-art methods focus mainly on the global transfer of the light and color distributions. Unfortunately, the use of a global distribution is questionable since the light and color of image can significantly vary within the same scene. In [27], we proposed a new method to deal with the limitations of existing methods. The proposed method aims at performing the partitions of the input and reference images

into Gaussian distributed clusters by considering the main style of input and reference images. From this clustering, several novel policies are defined for mapping the clusters of the input and reference images. To complete the style transfer, for each pair of corresponding clusters, we apply a parametric color transfer method (i.e. Monge-Kantorovitch transformation) and a local chromatic adaptation transform. Results, subjective user evaluation as well as objective evaluation show that the proposed method obtains visually pleasing and artifact-free images, respecting the reference style. Some results are illustrated in Figure 3 .



Figure 3. From left to right: input image, reference image and the result of the proposed method.

In [34], we extended the method presented in [27] to deal with a color transfer between two HDR images. One limitation of the two proposed methods is that we are still considering that the distributions of color and light follow a Gaussian law. We are currently investigating a more general approach by considering multivariate generalized Gaussian distribution.

7.2.2. Image guided inpainting

Participants: Christine Guillemot, Thomas Maugey.

Inpainting of images has been intensively studied in the past few years, especially for applications such as image restoration and editing [16]. Another application where inpainting techniques are useful is view synthesis, where holes are to be filled corresponding to areas that are no longer occluded. In the particular cases where one has access to ground truth images (like for example in multiview video coding where view synthesis is used for predicting the captured views from a reference one), auxiliary information can be generated to help inpainting, which leads to the concept of *guided inpainting*.

In [29], we have proposed a new auxiliary information that is used to refine the set of candidate patches for the hole filling step of the inpainting. Assuming that the patches of an image lie in a union of subspaces, *i.e.*, the images have different regions with different color textures, these patches are first clustered using a new recursive spectral clustering algorithm that extends existing sparse subspace clustering and replaces the sparse approximation by locally linear embedding, better suited for the inpainting context. Dictionaries are then built from these clusters and used for the hole filling process. However, the inpainting is not always able to "guess" in which cluster the patches of the hole belong to (especially around discontinuities). The auxiliary information that is built from the ground truth image may help to find the right cluster. We thus propose a new guided inpainting algorithm that forces the patch reconstruction to be done in one cluster only, if no auxiliary information is available, or in the cluster pointed by the auxiliary information, if it is available. Experiments (Fig. 4) show that auxiliary information helps to significantly improve the inpainting quality for a reasonable coding cost.

We are currently working on the extension of this technique in order to place the guided inpainting problem in an information theoretic framework, and better answer the following questions: when additional information is actually needed? What type of auxiliary information is needed? how to optimize in a rate-distortion sense the guided inpainting problem?.

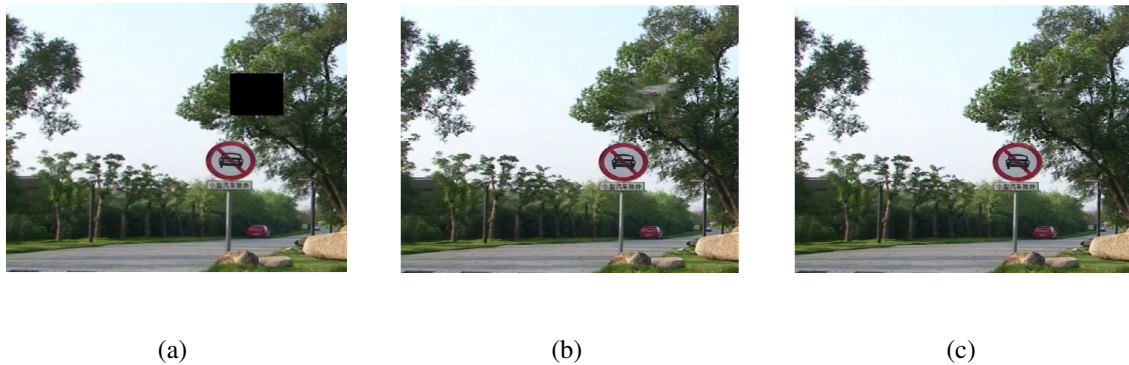


Figure 4. (a) input image to inpaint, (b) filled image using baseline not guided inpainting (c) filled image using proposed guided inpainting with an auxiliary information cost of 0.018 bpp bitrate.

7.2.3. Clustering on manifolds for image restoration

Participants: Julio Cesar Ferreira, Christine Guillemot, Elif Vural.

Local learning of sparse image models has proven to be very effective to solve a variety of inverse problems in many computer vision applications. To learn such models, the data samples are often clustered using the K-means algorithm with the Euclidean distance as a dissimilarity metric. However, the Euclidean distance may not always be a good dissimilarity measure for comparing data samples lying on a manifold. We have developed two algorithms for determining a local subset of training samples from which a good local model can be computed for reconstructing a given input test sample, where we take into account the underlying geometry of the data. The first algorithm, called Adaptive Geometry-driven Nearest Neighbor search (AGNN), is an adaptive scheme which can be seen as an out-of-sample extension of the replicator graph clustering method for local model learning. The second method, called Geometry-driven Overlapping Clusters (GOC), is a less complex nonadaptive alternative for training subset selection. The AGNN and GOC methods have been evaluated in image super-resolution, deblurring and denoising applications and shown to outperform spectral clustering, soft clustering, and geodesic distance based subset selection in most settings.

7.3. Representation and compression of large volumes of visual data

Sparse representations, data dimensionality reduction, compression, scalability, perceptual coding, rate-distortion theory

7.3.1. Manifold learning and low dimensional embedding for classification

Participants: Christine Guillemot, Elif Vural.

Typical supervised classifiers such as SVM are designed for generic data types and do not make any particular assumption about the geometric structure of data, while data samples have an intrinsically low-dimensional structure in many data analysis applications. Recently, many supervised manifold learning methods have been proposed in order to take the low-dimensional structure of data into account when learning a classifier. Unlike unsupervised manifold learning methods which only take the geometric structure of data samples into account when learning a low-dimensional representation, supervised manifold learning methods learn an embedding that not only preserves the manifold structure in each class, but also enhances the separation between different classes.

An important factor that influences the performance of classification is the separability of different classes in the computed embedding. We have done a theoretical analysis of separability of data representations given by supervised manifold learning. In particular, we have focused on the nonlinear supervised extensions of the Laplacian eigenmaps algorithm and have examined the linear separation between different classes in the learned embedding. We have shown that, if the graph is such that the inter-group graph weights are sufficiently small, the learned embedding becomes linearly separable at a dimension that is proportional to the number of groups. These theoretical findings have been confirmed by experimentation on synthetic data sets and image data.

We have then considered the problem of out-of-sample generalizations for manifold learning. Most manifold learning methods compute an embedding in a pointwise manner, i.e., data coordinates in the learned domain are computed only for the initially available training data. The generalization of the embedding to novel data samples is an important problem, especially in classification problems. Previous works for out-of-sample generalizations have been designed for unsupervised methods. We have studied this problem for the particular application of data classification and proposed an algorithm to compute a continuous function from the original data space to the low-dimensional space of embedding. In particular, we have constructed an interpolation function in the form of a radial basis function that maps input points as close as possible to their projections onto the manifolds of their own class. Experimental results have shown that the proposed method gives promising results in the classification of low-dimensional image data such as face images.

7.3.2. Adaptive clustering with Kohonen self-organizing maps for second-order prediction

Participants: Christine Guillemot, Bihong Huang.

The High Efficiency Video Coding standard (HEVC) supports a total of 35 intra prediction modes which aim at reducing spatial redundancy by exploiting pixel correlation within a local neighborhood. However the correlation remains in the residual signals of intra prediction, leading to some high energy prediction residuals. In 2014, we have studied several methods to exploit remaining correlation in residual domain after intra prediction. These methods are based on vector quantization with codebooks learned and dedicated to the different prediction modes in order to model the directional characteristics of the residual signals. The best matching code vector is found in a rate-distortion optimization sense. Finally, the index of the best matching code vector is sent to the decoder and the vector quantization error, the difference between the intra residual vector and the best matching code vector, is processed by the conventional operations of transform, scalar quantization and entropy coding.

In a first approach called MDVQ (Mode Dependent Vector Quantization), the codebooks were learned using the k-means algorithm [26]. More recently, we have developed a variant of the approach, called AMDVQ (Adaptive MDVQ) by adding a codebook update step based on Kohonen Self-Organized Maps which aims at capturing the variations of the residual signal statistical characteristics. The Kohonen algorithm uses previously reconstructed residual vectors to continuously update the code vectors during the encoding and decoding of the video sequence [12].

7.3.3. Rate-distortion optimized tone curves for HDR video compression

Participants: David Gommelet, Christine Guillemot, Aline Roumy.

High Dynamic Range (HDR) images contain more intensity levels than traditional image formats. Instead of 8 or 10 bit integers, floating point values requiring much higher precision are used to represent the pixel data. These data thus need specific compression algorithms. In collaboration with Envivio, we have developed a novel compression algorithm that allows compatibility with the existing Low Dynamic Range (LDR) broadcast architecture in terms of display, compression algorithm and data rate, while delivering full HDR data to the users equipped with HDR display. The developed algorithm is thus a scalable video compression offering a base layer that corresponds to the LDR data and an enhancement layer, which together with the base layer corresponds to the HDR data. The novelty of the approach relies on the optimization of a mapping called Tone Mapping Operator (TMO) that maps efficiently the HDR data to the LDR data. The optimization has been carried out in a rate-distortion sense: the distortion of the HDR data is minimized under the constraint of

minimum sum datarate (for the base and enhancement layer), while offering LDR data that are close to some “aesthetic” a priori. Taking into account the aesthetic of the scene in video compression is novel, since video compression is traditionally optimized to deliver the smallest distortion with the input data at the minimum datarate.

7.3.4. Local Inverse Tone Curve Learning for HDR Image Scalable Compression

Participants: Christine Guillemot, Mikael Le Pendu.

In collaboration with Technicolor, we have developed local inverse tone mapping operators for scalable high dynamic range (HDR) image coding. The base layer is a low dynamic range (LDR) version of the image that may have been generated by an arbitrary Tone Mapping Operator (TMO). No restriction is imposed on the TMO, which can be either global or local, so as to fully respect the artistic intent of the producer. The method which has been developed successfully handles the case of complex local TMOs thanks to a block-wise and non-linear approach [28]. A novel template based Inter Layer Prediction (ILP) is designed in order to perform the inverse tone mapping of a block without the need to transmit any additional parameter to the decoder. This method enables the use of a more accurate inverse tone mapping model than the simple linear regression commonly used for blockwise ILP [21]. In addition, this paper shows that a linear adjustment of the initially predicted block can further improve the overall coding performance by using an efficient encoding scheme of the scaling parameters. Our experiments have shown an average bitrate saving of 47% on the HDR enhancement layer, compared to previous local ILP methods.

7.3.5. HEVC-based UHD video coding optimization

Participants: Nicolas Dhollande, Christine Guillemot, Olivier Le Meur.

The HEVC (High Efficiency Video Coding) standard brings the necessary quality versus rate performance for efficient transmission of Ultra High Definition formats (UHD). However, one of the remaining barriers to its adoption for UHD content is the high encoding complexity. We address the reduction of HEVC encoding complexity by investigating different strategies: First we have proposed to infer UHD coding modes and quad-tree from a first encoding pass which consists in encoding a lower resolution version of the input video. In the context of our study, the first encoding pass encodes a HD video sequence. A speed-up by a factor of 3 is achieved compared to directly encoding the UHD format without compromising the final video quality. The second strategy focuses on the block partitioning of intra frame coding. The Coding Tree Unit (CTU) is the root of the coding tree and can be recursively split into four square Coding Unit (CU), given that the smallest block size is 8×8 . Once the partitioning procedure is fully completed, the final quad-tree can be obtained by choosing the configuration leading to the best rate-distortion trade-off. Rather than performing an exhaustive partitioning, we aim to predict the quad-tree partition into coding units (CU). This prediction is based on low-level visual features extracted from the video sequences. The low-level features are related to gradient-based statistics, structure tensors statistics or entropy etc. From these features, we trained a probabilistic model on a set of UHD training sequences in order to determine whether the coding unit should be further split or not. The proposed methods yield a significant encoder speed-up ratio (up to 5.3 times faster) with a moderate loss in terms of compression efficiency [33].

7.4. Distributed processing and robust communication

Information theory, stochastic modelling, robust detection, maximum likelihood estimation, generalized likelihood ratio test, error and erasure resilient coding and decoding, multiple description coding, Slepian-Wolf coding, Wyner-Ziv coding, information theory, MAC channels

7.4.1. Information theoretical bounds of Free-viewpoint TV

Participants: Thomas Maugey, Aline Roumy.

Free-viewpoint television FTV is a new system for viewing video where the user can choose its viewpoint freely and change it at anytime. The goal is to propose an immersive sensation without the disadvantage of Three-dimensional (3D) television (TV). Indeed, the conventional 3D displays (with or without glasses) occur, by construction, an accommodation-vergence conflict: since the eye tend to focus on the display screen (accommodation), whereas the brain perceives the depth of 3D images due to the different views seen by each eye (vergence). Instead, with FTV, a look-around effect is produced without any visual fatigue since the displayed images remain 2D. Therefore, FTV presents nice properties that makes it a serious competitor for 3DTV. Existing compression algorithms for FTV consider to send all the views, which would require about 100 Mbits/s (for 100 views, as needed to propose a true navigation within the scene). Since this amount does not fit the current datarate for transmission in a streaming scenario, we investigate a solution where the server only send the request. In [31], [30], we have shown a very surprising and positive result: if all the views are compressed once and if the server extracts from the compressed bitstream the request (i.e. one view at a time), the datarate is exactly the same as if the whole database was entirely decoded, and the requested views reencoded. This very positive result shows that it is possible to send FTV with the same datarate as single view television with very limited computational cost at the server (only extraction from the bistream). This result is an information theoretical result and the goal is now to build a practical system that can achieve this performance.

7.4.2. Compressed Sensing : a probabilistic analysis of the orthogonal matching pursuit algorithm

Participant: Aline Roumy.

Compressed sensing (CS) is an efficient acquisition scheme, where the data are projected onto a randomly chosen subspace to achieve data dimensionality reduction. The projected data are called measurements. The reconstruction is performed from these measurements, by solving underdetermined linear systems under a sparsity a priori constraint. It is generally believed that the greedy algorithm Orthogonal Matching pursuit performs well and can determine which variables are active (i.e. non zero). In contrast, we showed that this is not the case even in the noiseless context. We derived an exact probabilistic analysis of the iterative algorithm in the large system regime, when all dimensions tend to infinity. We showed that as the number of iterations grows, the algorithm will make errors with probability one.

SMIS Project-Team

6. New Results

6.1. Embedded Data Management

Participants: Nicolas Ancaux, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa [correspondent].

Embedded keyword indexing: In this work, we revisit the traditional problem of information retrieval queries over large collections of files in an embedded context. A file can be any form of document, picture or data stream, associated with a set of terms. A query can be any form of keyword search using a ranking function (e.g., TF-IDF) identifying the top-k most relevant files. The proposed search engine can be used in sensors to search for relevant objects in their surroundings, in cameras to search pictures by using tags, in personal smart dongles to secure the querying of documents and files hosted in an untrusted Cloud, or in a personal cloud securely managed using a tamper resistant smart object. A search engine is usually based on a (large) inverted index and queries are traditionally evaluated by allocating one container in RAM per document to aggregate its score, making the RAM consumption linear with the size of the document corpus. To tackle this issue, we designed a new form of inverted index which can be accessed in a pure pipeline manner to evaluate search queries without materializing any intermediate result. Successive index partitions are written once in Flash and maintained in the background by timely triggering merge operations while files are inserted or deleted from the index. By combining this new index and the corresponding evaluation techniques, our embedded search engine is capable of reconciling high insert/delete/update rate and query scalability. We have demonstrated the search engine on a secure USB token in the context of a personal cloud, and have conducted in depth performance evaluations on a development board representative for different smart objects characteristics. The experimental results demonstrate the scalability of the approach and its superiority compared to state of the art methods. This work was published at VLDB'15 [21] and demonstrated at SIGMOD'15 [24]. It constitutes the main contribution of the PhD thesis of Saliha Lallali

Spatio-temporal indexing in Flash storage: The convergence of mobile computing, wireless communications and sensors has raised the development of many applications exploiting massive flows of spatio-temporal data such as in location-based services, participatory sensing, or traffic management [15]. Spatio-temporal data indexing is among the most active research topics in this area. Nevertheless, since a few years a new fundamental parameter has made its entry on the database scene: the NAND flash storage. The peculiar characteristics of flash memory require redesigning the existing data storage and indexing techniques that were devised for magnetic hard-disks. TRIFL, proposed in [16] is an efficient and generic TRajjectory Index for FLash, designed around the key requirements of both trajectory indexing and flash storage. TRIFL is generic in the sense that it is efficient for both simple flash storage devices such as the SD cards and more powerful devices such as the solid state drives. In addition, TRIFL includes an online self tuning algorithm that allows adapting the index structure to the workload and the technical specifications of the flash storage device to maximize the index performance. Moreover, TRIFL achieves good performance with relatively low memory requirements, making it appropriate for many application scenarios. The experimental evaluation shows that TRIFL outperforms the representative indexing methods on flash disks but also on magnetic disks. This work [15] [16] is part of Dai Hai Ton That's Ph.D. thesis, co-supervised by Iulian Sandu Popa.

6.2. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Philippe Pucheral, Quoc Cuong To.

Asymmetric Architecture Computing: This research direction studies the secure execution of various algorithms on data stored in an unstructured network of Trusted Cells (i.e., personal trusted device) so that each user can keep control over her data. The data could be stored locally in a trusted cell or encrypted on some external cloud. Execution takes place on a specific infrastructure called the Asymmetric Architecture: the network of trusted cells, supported by an untrusted cloud supporting IaaS or PaaS. Our objective is to show that many different algorithms and computing paradigms can be executed on the Asymmetric Architecture, thus achieving secure and private computation. Our first contribution in this area was to study the execution of Privacy Preserving Data Publishing (PPDP) algorithms on such an architecture, and provided generic protocols to deal with all kinds of PPDP algorithms, which are robust against honest-but-curious and malicious adversaries [2][3]. Our second contribution was to study general SQL queries in this same execution context. We concentrated on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers [9]. This work was part of Quoc-Cuong To's Ph.D defended in 2015 [13]. We are extending this general framework through a collaboration with INSA Centre Val de Loire, LIFO Lab and University of Paris Nord, LIPN lab, to study the secure execution of Map/Reduce on the Asymmetric Architecture. Computing MapReduce processes on the Asymmetric Architecture means maintaining the flexibility and efficiency of MapReduce, while adding security into the mix. We have shown in [25] that it is possible to achieve seamless integration of distributed MapReduce processing using trusted cells, while maintaining reasonable performance.

Secure spatio-temporal distributed processing: Mobile participatory sensing could be used in many applications such as vehicular traffic monitoring, pollution tracking, or even health surveying (e.g., to allow measuring in real-time the individual exposure to environmental risk factors or the propagation of an epidemic). However, its success depends on finding a solution for querying a large number of users which protects user location privacy and works in real-time. We addressed these issues and proposed PAMPAS, a privacy-aware mobile distributed system for efficient data aggregation in mobile participatory sensing. In PAMPAS, mobile devices enhanced with secure hardware, called secure probes, perform distributed query processing, while preventing users from accessing other users' data. Secure probes exchange data in encrypted form with help from an untrusted supporting server infrastructure. PAMPAS uses two efficient, parallel, and privacy-aware protocols for location-based aggregation and adaptive spatial partitioning of secure probes. Our experimental results and security analysis demonstrate that these protocols are able to collect, aggregate and share statistics or derived data in real-time, without any privacy leakage. This work is part of Dai Hai Ton That's Ph.D. thesis, co-supervised by Iulian Sandu Popa. The system implementation was demonstrated in [26], and a paper describes the technical details of the system [31].

6.3. Personal Cloud

Participants: Nicolas AnCIAUX [correspondent], Luc Bouganim, Athanasia Katsouraki, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Paul Tran Van.

We are witnessing an exponential increase in the acquisition of personal data about the individuals or produced by them. Today, this information is managed using Web applications, centralizing this data in cloud data servers, under the control of few Web majors [4]. However, it has now become clear that (1) centralizing millions of personal records exposes the data to very sophisticated attacks, linked to a very high potential benefit in case of success (millions of records being revealed), and (2) delegating the management of personal records without any tangible guarantee for the individuals leads to privacy violations, the data being potentially made accessible to other organizations (e.g., governments, commercial partners) and being subject to lucrative secondary usages (not advertised to the individuals). To face this situation, many recent initiatives push towards the emergence of the Personal Cloud paradigm. A personal cloud can be viewed as a personal server, owned by a given individual, which gives to its owner the ability to store her complete digital environment, synchronize it among various devices and share it with other individuals and applications under control. In the SMIS team, we claim the need of a Secure Personal Cloud, and promote the introduction of a secure (tamper resistant) data engine in the architecture [1]. On this basis, we investigate new data sharing and dissemination models, where usage and access control rules endorsed by the individuals could be enforced and have presented this

vision at EDBT'14 and at ADBIS'15 [18]. We have started a cooperation with the startup CozyCloud at the end of 2014. A contract was signed at the end of 2014 to integrate PlugDB in a CozyCloud instance and the PhD of Paul Tran Van (CIFRE SMIS-CozyCloud) has started to explore new data sharing techniques which could be enforced in the secure personal cloud model. A second PhD CIFRE SMIS-CozyCloud is being submitted to explore privacy-preserving distributed computations over personal clouds. Athanasia Katsouraki is working on privacy issues and on adoption of the secure data engine [29] in cooperation with the economists (CERDI) in the context of the Digital Society Institute (DSI). A paper written by jurists, economists and computer scientists from DSI has been invited for publication in Legicom'2016 to present our common vision of Privacy-by-Design principles in the context of Open Data and Internet of Things.

6.4. Applications

Participants: Nicolas Ancaux [correspondent], Luc Bouganim, Philippe Pucheral.

In 2014, we proposed a new paradigm, that we call Folk-enabled Information System (Folk-IS), based on a fully decentralized and participatory approach, where each individual implements a small subset of a complete information system without the need for a shared networked infrastructure [5]. Folk-IS builds upon the emergence of highly secure, portable and low-cost storage and computing devices, called hereafter Smart Tokens. Here however, the focus is on low-cost of ownership, deployment and maintenance, and on the absence of a networked infrastructure. With Folk-IS and thanks to their smart tokens, people will transparently and opportunistically perform data management and networking tasks as they physically move, so that IT services are truly delivered by the crowd. Following this work, we collaborate with researchers and doctors from Cameroon to study the specific case of diabetes follow-up. Indeed, there are currently more than half a million diabetes cases in Cameroon and the deaths caused by diabetes complications will double before 2030. Diabetes complications mostly occur due to a bad follow-up of patients. Based on an analysis of the current situation, we proposed a new IT architecture for diabetes follow-up and introduce the bases of a new distributed computation protocol for this architecture. Our approach does not require any preexisting support communication infrastructure, can be deployed at low cost, and provides strong privacy and security guarantees. This work, published in AFRICOM [20] envisions an experiment in the field we plan to conduct under the authority of the Cameroonian National Center for Diabetes and Hypertension, with a potential for generalization to other diseases.

STARS Project-Team

7. New Results

7.1. Introduction

This year Stars has proposed new results related to its three main research axes : perception for activity recognition, semantic activity recognition and software engineering for activity recognition.

7.1.1. Perception for Activity Recognition

Participants: Julien Badie, Slawomir Bak, Piotr Bilinski, François Brémond, Duc Phu Chau, Etienne Corvée, Antitza Dancheva, Kanishka Nithin Dhandapani, Carolina Garate, Furqan Muhammad Khan, Michal Koperski, Thi Lan Anh Nguyen, Javier Ortiz, Ujjwal Ujjwal.

The new results for perception for activity recognition are:

- Pedestrian Detection using Convolutional Neural Networks (see 7.2)
- Head detection for eye tracking application (see 7.3)
- Minimizing hallucination in Histogram of Oriented Gradients (see 7.4)
- Hybrid approaches for Gender estimation (see 7.5)
- Automated Healthcare: Facial-expression-analysis for Alzheimer's patients in musical mnemotherapy (see 7.6)
- Robust Global Tracker based on an Online Estimation of Tracklet Descriptor Reliability (see 7.7)
- Optimizing people tracking for a video-camera network (see 7.8)
- Multi-camera Multi-object Tracking and Trajectory Fusion (see 7.9)
- Person Re-Identification in Real-World Surveillance Systems (see 7.10)
- Human Action Recognition in Videos (see 7.11)

7.1.2. Semantic Activity Recognition

Participants: Vasanth Bathrinarayanan, François Brémond, Duc Phu Chau, Serhan Cosar, Alvaro Gomez Uria Covella, Carlos Fernando Crispim Junior, Ramiro Leandro Diaz, Giuseppe Donatiello, Baptiste Fosty, Carolina Garate, Alexandra Koenig, Michal Koperski, Farhood Negin, Thanh Hung Nguyen, Min Kue Phan Tran, Philippe Robert.

For this research axis, the contributions are :

- Evaluation of Event Recognition without using Ground Truth (see 7.12)
- Semantic Event Fusion of Different Visual Modality Concepts for Activity Recognition (see 7.13)
- Semi-supervised activity recognition using high-order temporal-composite patterns of visual concepts (see 7.14)
- From activity recognition to the assessment of seniors' autonomy (see 7.15)
- Serious Games Interfaces using an RGB-D camera (see 7.16)
- Assistance for Older Adults in Serious Game using an Interactive System (see 7.17)
- Generating Unsupervised Models for Online Long-Term Daily Living Activity Recognition (see 7.18)

7.1.3. Software Engineering for Activity Recognition

Participants: Sabine Moisan, Annie Ressouche, Jean-Paul Rigault, Ines Sarray, Imane Khalis, Nazli Temur, Daniel Gaffé, Rachid Guerchouche, Matias Marin, Etienne Corvée, Carolina Da Silva Gomes Crispim, Anais Ducoffe, Jean Yves Tigli, François Brémond.

The contributions for this research axis are:

- Run-time Adaptation of Video Systems (see 7.19)
- Scenario Description Language (see 7.20)
- Scenario Recognition (see 7.21)
- The Clem Workflow (see 7.22)
- Safe Composition in WComp Middleware for Internet of Things (see 7.23)
- Design of UHD panoramic video camera (see 7.24)
- Brick & Mortar Cookies (see 7.25)
- Monitoring Older People Experiments (see 7.26)

7.2. Pedestrian Detection using Convolutional Neural Networks

Participants: Ujjwal Ujjwal, François Brémond.

Keywords: Pedestrian detection, CNN

The objective of the work was to perform pedestrian detection in different settings. The settings corresponded to different types of camera-views as well as different types of camera settings (e.g- moving camera vs. static camera). The work followed a wide range of experiments using different public implementations of convolutional neural networks and on different types of datasets. We detail the experiments one by one in the following subsections :

Experiments on CNN architectures

We started with an evaluation of different CNN architectures for pedestrian detection. Towards this end, we implemented three important and famous architectures - LeNet [72], AlexNet [69] and CifarNet [68]. For the purpose of training and validation we extracted patches from the public datasets of Inria [55], Daimler [58], TUD-Brussels [92], Caltech [57], ViPer [62], USC [93] and MIT [78]. The breakup of the dataset used for training was as shown in table 1 .

Table 1. Training, testing and validation patch details for CNN training

	Pedestrian Patches	Non-Pedestrian Patches
Training	131,183	61,500
Validation	65,591	30,700
Testing	65,591	30,600

Implementation of all the three models for pedestrian detection which gave a very high accuracy (94.2% (LeNet), 98% (AlexNet) and 98.2% (CifarNet)) for classification at patch level. Though these results were good at the patch level, more thorough understanding was needed to determine the effect of network architecture on classification. This was important because the three architectures vary greatly in terms of number of layers and other parameters. Moreover the practical problem in pedestrian detection chiefly involves detecting pedestrians in an image (i.e. when full-scale images instead of pre-defined patches are available).

The first set of experiments was done using sliding windows. This had to be abandoned soon, since for each image this was taking an impractical time ($> 3minutes/image$). Moreover sliding window is less suited in its naive setting due to the fact that each candidate window had to be rescaled to meet the network input size and tested individually by extracting features over it. This was followed by efforts to understand and implement a wide range of other techniques for full-scale detection using CNN. This is still an open problem though some encouraging advancements through R-CNN [61] and OverFeat [83] have been made. A major difficulty lies in lack of robust implementations of CNN which allow for integrated training and testing with object localization. Moreover existing implementations are less flexible and often make it difficult to carry out modifications required to implement new techniques independently.

We settled with the R-CNN which uses region proposals extracted using selective search [87] to extract object proposals and then train a CNN using those proposals and subsequently classifying using a SoftMax classifier or a SVM.

The evaluation was done on both moving cameras and static cameras and the evaluation showed that the network was performing a little satisfactorily, though below the state-of-art performance standards. The performance metric was Average Miss Rate (AMR) Vs. False Positives Per Image (FP/I). A good detector must exhibit a very low AMR alongwith very low FP/I. Table 2 summarizes the detection results, with table 3 summarizing state-of-art results on different pedestrian detection datasets.

Table 2. Our R-CNN results on different pedestrian datasets.

Dataset	#Images	AMR	FP/I
Inria	741	0.27	0.36
DAIMLER	15K	0.38	0.40
Caltech	16K	0.46	0.43
USC	584	0.02	0
PETS 2009 S2.L1	5565	0.42	0.29
PETS 2009 S2.L2	1744	0.35	0.19

Table 3. State-of-art results on different pedestrian datasets.

Dataset	#Images	AMR	FP/I
Inria	741	0.14	0.1
DAIMLER	15K	0.29	0.21
Caltech	16K	0.12	0.1
PETS 2009 S2.L1	5565	0.22	0.1
PETS 2009 S2.L2	1744		

It was felt that more work is needed to organize a CNN library and subsequently work to improve the above results.

Conclusion

Pedestrian detection finds its applications in different settings. It is also highly influenced by a wide variety of variations which have many practical ramifications in areas such as surveillance. It is important to develop a robust and high-performance system for pedestrian detection that is able to take into account a very wide range of such variations such as occlusion and poor visibility. CNNs have shown great promise in object detection and recognition lately and this inspires its growing applications in pedestrian detection. While the current results of our R-CNN experiments do not match the state-of-art it has shown some promise by providing consistent numbers across datasets which shows that CNNs are a good way to transcend a system beyond dataset-specific restrictions. An important factor is the instance of moving cameras vis-a-vis static cameras. While the present experiments show that decent performance is obtained on moving camera databases, consistent and similar performance is also obtained in the context of static camera databases such as PETS. This shows that with better training and improved practices of dataset handling such as augmentation and dataset structuring by clustering based methods can help in pushing the performance to acceptable levels for applications in automated surveillance and driving applications.

Further Work We intend to take this study forward, by looking into novel approaches to gather more information about a pedestrian dataset from CNN, while further increasing the detection results.

7.3. Head Detection for Eye Tracking Application

Participants: Thanh Hung Nguyen, Antitza Dantcheva, François Brémond.

Keywords: computer vision, head detection, eye tracking

Head detection [77] uses RGBD sensor (Kinect 2 sensor) which is supported by SUP platform of STARS team. For the eye tracking, we use the open-source library (OpenBR) which has good performance in our test.

Until now, the head detector was working well when people were standing but not good enough when people were lying or bending as you can see on Figure 5 and Figure 6. This experiment was realised on simple datasets where mostly people was close to camera and walking. The main reason is the lack of samples in the learning process for the challenges cases (lying, bending). So at this moment we are collecting the head samples for it.

Number of true positives	209
Number of false positives	28
Number of false negatives	25

Figure 5. Performance of head detection on simple dataset

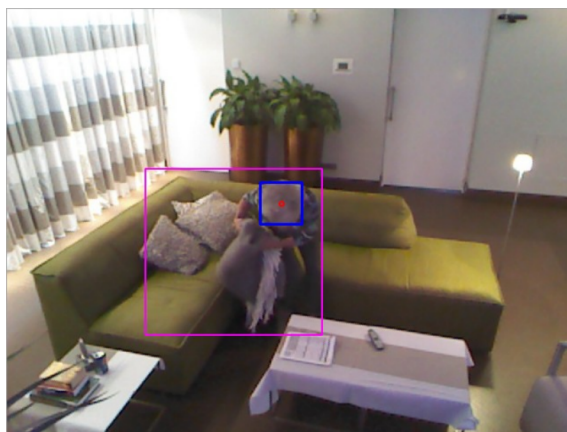


Figure 6. An example of experiment

7.4. Minimizing Hallucination in Histogram of Oriented Gradients

Participants: Javier Ortiz, Slawomir Bak, Michal Koperski, François Brémond.

keywords: computer vision, action recognition, re-identification,

Challenges in histogram representation

In computer vision very often histogram of values is used as a feature representation. For instance HOG descriptor is in fact histogram of gradient orientations. Also histogram of codewords in Bag of Words representation is a very popular action recognition representation. In such a case we are not interested in absolute values for given bin, but rather in shape of histogram to find the patterns. To make histogram representation independent from absolute values we use L1 or L2 normalization. In the normalization process we treat histogram as a vector and we transform it in the way that it should have unit length according to given norm (L1, L2). The drawback of such approach is that normalization process may amplify the noise for histograms where absolute values are very small (no pattern or only noise). Such histogram after normalization can be very similar to histogram with strong absolute values. This situation is showed in figure 7. Although original histogram in second row of figure 7 contains almost no information after normalization the values are amplified and the result is exactly the same as for histogram in the first row. Such behavior is called hallucinations.

Proposed normalization method

We propose to add an artificial bin with given value which would prevent small noisy values from being amplified during normalization process. In figure 7 in first row we show artificial bin in pink color. Thanks to that histograms after normalization (last column) are different. If we analyze cumulative sum of histogram values across data (we sum values of all bins in whole histogram and we draw distribution), we can find that in some data-sets we obtain bi-nomial distribution see figure 8. The gap between two Gaussians indicates convenient border between noise and meaningful data. On the other hand many data-sets do not have this bi-nomial feature and for that we do not have formalization to find the value of artificial bin. This problem would be a subject of further studies.

The following method was successfully applied to person re-identification and action recognition problem. Further details can be found in the paper [42]

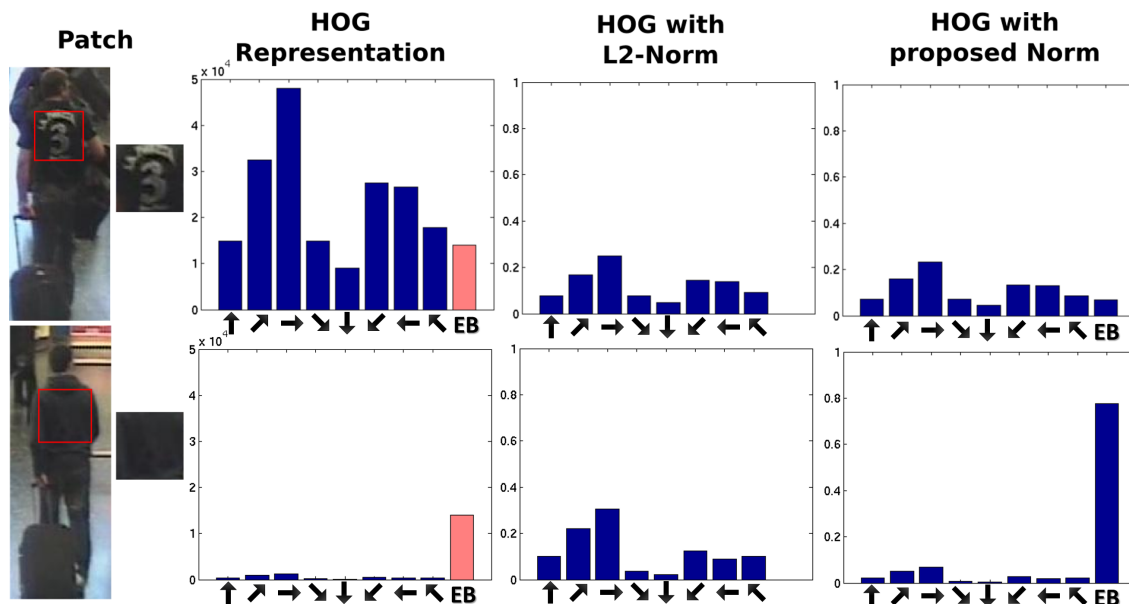


Figure 7. HOG representations of two patches with different amount of texture. Each bin in blue represents the sum of magnitudes of edges for a particular orientation. The bin marked as EB represents the proposed extra bin

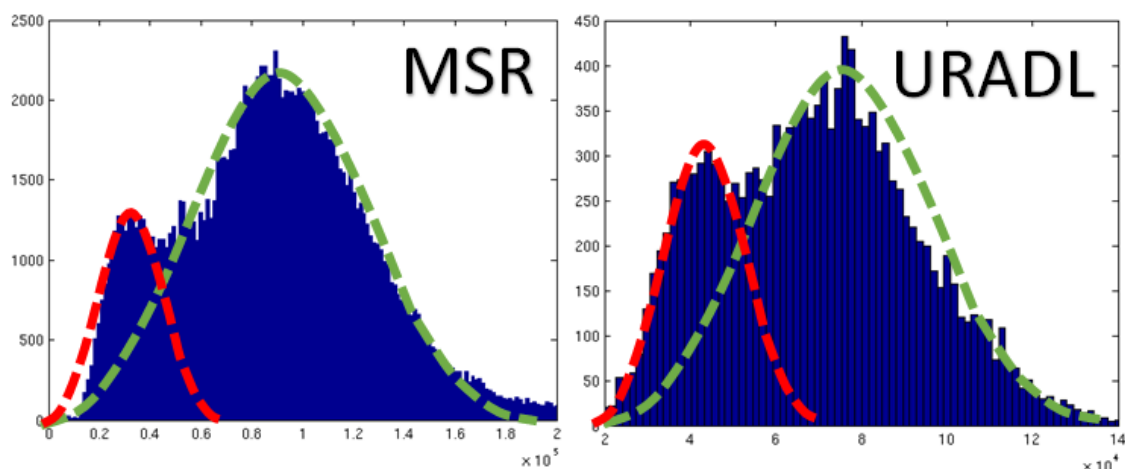


Figure 8. Distribution of the sum of bins for randomly sampled patches for URADL and MSRDailyActivity3D data-set.

7.5. Hybrid Approaches for Gender estimation

Participants: Antitza Dantcheva, François Brémond, Philippe Robert.

keywords: gender estimation, soft biometrics, biometrics, visual attributes

Automated gender estimation has numerous applications including video surveillance, human computer-interaction, anonymous customized advertisement and image retrieval. Gender estimation remains a challenging task, which is inherently associated with different biometric modalities including fingerprint, face, iris, voice, body shape, gait, as well as clothing, hair, jewelry and even body temperature [31]. Recent work has sought to further the gains of single-modal approaches by combining them, resulting into hybrid cues that offer a more comprehensive gender analysis, as well as higher resilience to degradation of any of the single sources.

Can a smile reveal your gender?

In this work we propose a novel method for gender estimation, namely the use of dynamic features gleaned from smiles and show that (a) facial dynamics can be used to improve appearance-based gender-estimation, (b) that while for adults appearance features are more accurate than dynamic features, for subjects under 18 years old facial dynamics outperform appearance features. While it is known that sexual dimorphism concerning facial appearance is not pronounced in infants and teenagers, it is interesting to see that facial dynamics provide already related clues. The proposed system, fusing a state-of-the-art appearance and dynamics-based features (see Figure 9), improves the appearance based algorithm from 78.0% to 80.8% in video-sequences of spontaneous smiles and from 80% to 83.1% in video-sequences of posed smiles for subjects above 18 years old (see Table 4). These results show that smile-dynamics include pertinent and complementary information to appearance gender information.

While this work studies video sequences capturing frontal faces expressing human smiles, we can envision that additional dynamics, such as other facial expressions or head and body movements carry gender information as well.

Distance-based gender prediction: What works in different surveillance scenarios?

Table 4. True gender classification rates. Age given in years.

Age	< 10	10 – 19	20 – 29	30 – 39	40 – 49	> 49
Subj. amount	48	95	60	49	72	33
OpenBR	58.33%	50.53%	81.67%	75.51%	75%	81.82%
Combined Age-Groups	< 20		> 19			
Subj. amount	143		214			
OpenBR	52.45%		78.04%			
Dynamics (SVM+PCA)	60.1%		69.2%			
Dynamics (AdaBoost)	59.4%		61.7%			
OpenBR + Dynamics (Bagged Trees)	60.8%		80.8%			

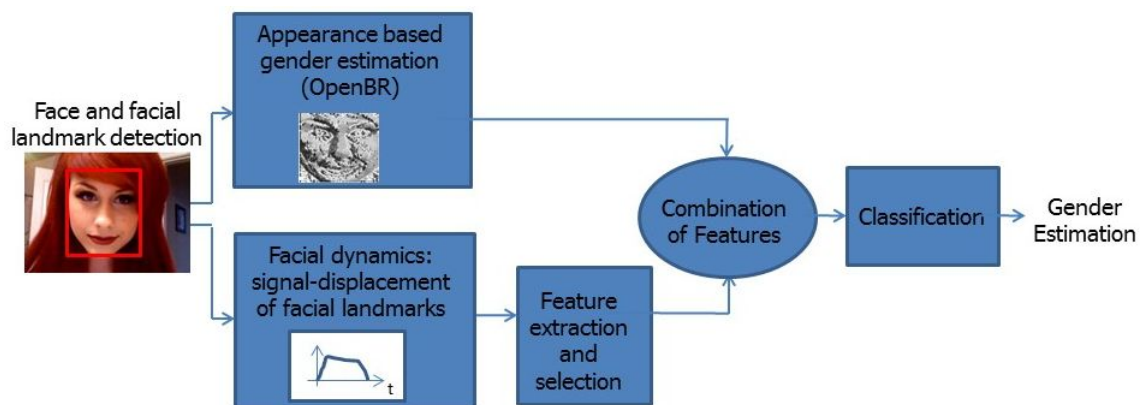


Figure 9. General Scheme of the facial appearance and dynamics framework.

In this work we fuse features extracted from face, as well as from body silhouette towards gender estimation. Specifically, for face, a set of facial features from the OpenBR library, including histograms of local binary pattern (LBP) and scale-invariant feature transform (SIFT) are computed on a dense grid of patches. Subsequently, the histograms from each patch are projected onto a subspace generated using PCA in order to obtain a feature vector, followed by a Support Vector Machine (SVM) used for the final face-based-gender decision. Body based features include geometric and color based features, extracted from body silhouettes, obtained by background subtraction, height normalization and SVM-classification for the final body-based-gender-decision. We present experiments on images extracted from video-sequences, emphasizing on three distance-based settings: close, medium and far from the TunnelDB dataset (see Figure 10). As expected, while face-based gender estimation performs best in the close-distance-scenario, body-based gender estimation performs best when the full body is visible - in the far-distance-scenario (see Table 5). A decision-level-fusion of face and body-based features channels the benefits of both approaches into a hybrid approach, providing results that demonstrate that our hybrid approach outperforms the individual modalities for the settings medium and close.

Table 5. True gender classification rates. Age given in years.

System Distance	FAR	MEDIUM	CLOSE
FBGE	57.14	79.29	89.29
BBGE 89.3 85 79.3			
Fusion BBGE & FBGE	82.9	88.6	95



Figure 10. Three distance-based settings in the TunnelDB dataset: far, medium and close.

While the dataset is relatively unconstrained in terms of illumination, body and face are captured facing relatively frontally towards the camera. Future work will involve less constraints also towards the pose of humans.

7.6. Automated Healthcare: Facial-expression-analysis for Alzheimer's Patients in Musical Mnemotherapy

Participants: Antitza Dantcheva, François Brémond, Philippe Robert.

keywords: automated healthcare, healthcare monitoring, expression recognition

In this work we seek to apply computer vision towards increasing the life quality of patients with Alzheimer's disease (AD), and particularly in applying computer vision towards interventions to delay functional decline and to decrease the burden of the most disturbing behavioral symptoms. Towards this we design a smart interaction tool, that "reads" emotions of AD patients. This approach is becoming necessary now, because the increasing prevalence of chronic disorders and its impact on functional decline is challenging the sustainability

of healthcare systems. Firstly, we have assembled a dataset of video-sequences acquired in the Alzheimer's Disease - clinique Fondation GSF Noisiez. Multiple patients and sessions have been captured during musical mnemotherapy. We then have annotated several sequences per one of four facial expressions, that occur in the recorded dataset including: neutral, talking, smile and sad. We then proceed to classify these expressions for 10 patients based on two approaches, that we study individually, as well as fused. The first approach contains face detection, facial landmark localization and signal displacement analysis for different facial landmarks, which are ranked based on categorization-pertinence, fused and classified into one of the four expression-categories (see Figure 11). In the second approach, we use face detection, eyes-detection, face normalization and HOG-features, which we classify into one of the four expression-categories.

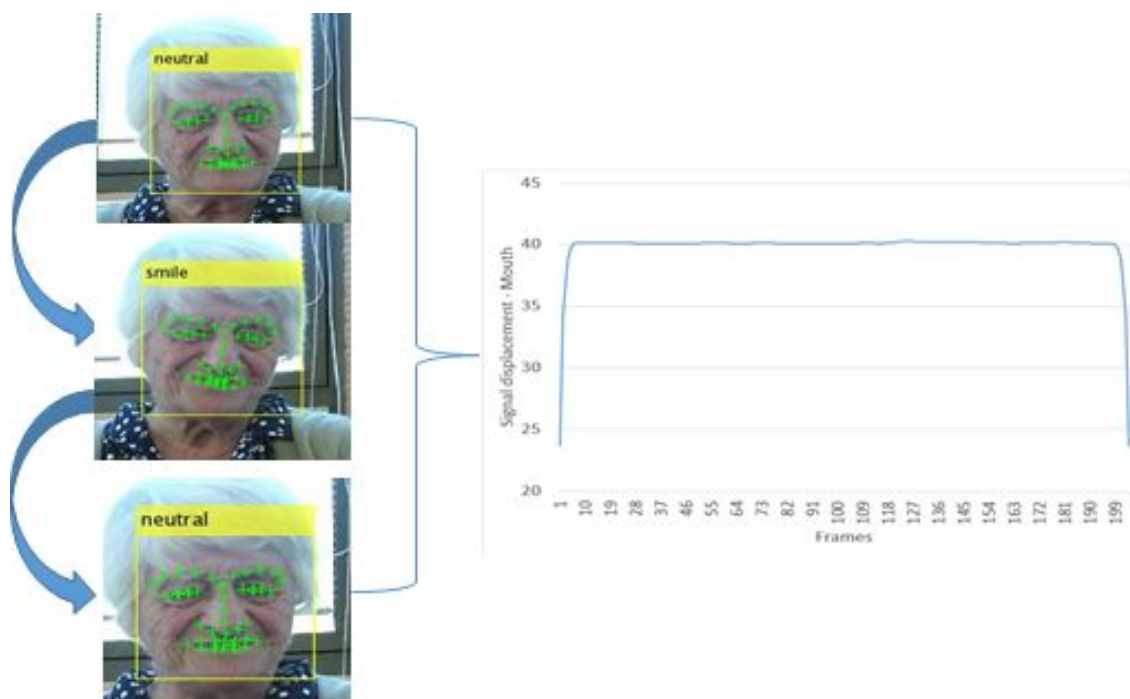


Figure 11. Smile detection based on signal displacement in the mouth region in the dataset collected at the Fondation GSF Noisiez in Biot, France.

The here used real-world-data challenges, as expected, all utilized computer vision algorithms (from face detection - due to no constraints of pose and illumination, to classifiers - due to a large intra-class-variation of facial expressions). Nevertheless, we obtain promising results that we envision improving by analyzing 2D and depth, as well as infrared data.

7.7. Robust Global Tracker Based on an Online Estimation of Tracklet Descriptor Reliability

Participants: Thi Lan Anh Nguyen, Chau Duc Phu, François Brémond.

Keywords: Tracklet fusion, Multi-object tracking

Object tracking - the process of locating a moving object (or multiple objects) in one camera or in a camera network over time - is an important part in surveillance video processing. However, the video context variation requires trackers to face plenty of challenges. For example, objects change their movement direction or their appearances, poses; objects are occluded by other objects or background; illumination is changed... In order to overcome above challenges, calculating the object appearance model overtime to adapt tracker to context variation is a necessary work.

In the state of the art, some online learning approaches [52], [48] have been proposed to track objects in various video scenes in each frame. These approaches learn online discriminative object descriptors to the current background as in [52] or learn an object appearance model which discriminates objects overtime as in [48]. However, the limitation of these approaches is that the reliability of object descriptor computed on only current frame is sensitive. False positives can reduce tracking quality. Furthermore, these algorithms try to find the discriminative descriptors or signatures of one object compared to its neighborhood but not considering the correlation of this object with its can-match candidates. Meanwhile, global tracking methods [91], [98] show their dominant ability over previous methods in noisy filtering. The approach in [98] proposes an algorithm that recovers fragmentation of object trajectories by using enhanced covariance-based signatures and an online threshold learning. The approach in [91] proposes a hierarchical relation hypergraph based tracker. These global tracking algorithms have significant results in matching short trajectories and filtering some noise. However, object descriptor weights are fixed for the whole video. Therefore, their tracking performances can be reduced if the scene changes.

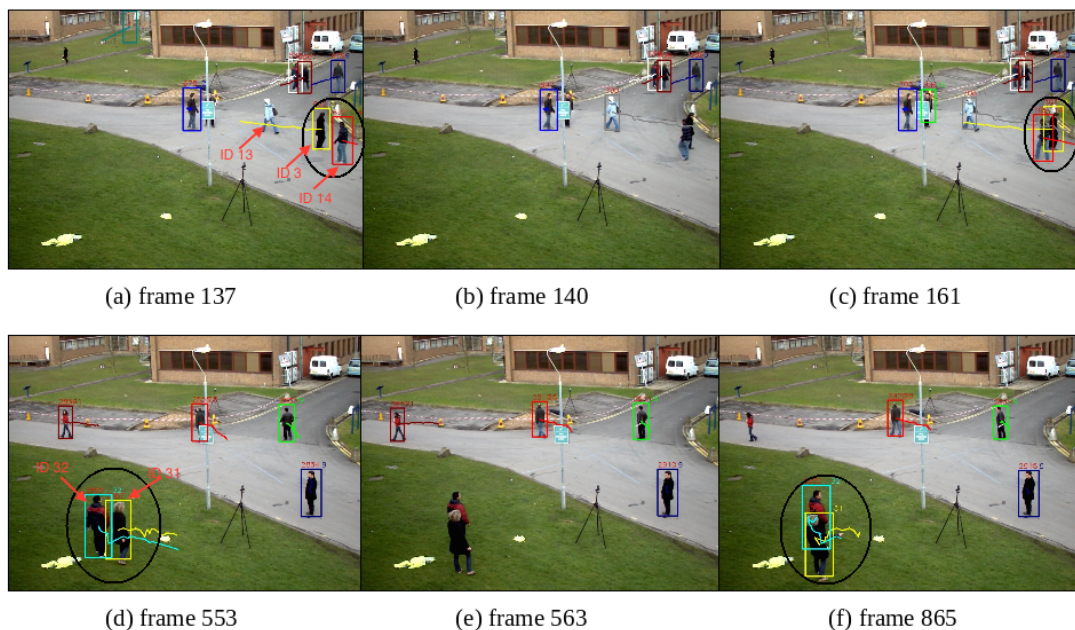


Figure 12. PETs2009 dataset: The online computation of discriminative descriptor weights depending on each video scene.

In this work, we propose a new approach to improve the tracking quality by a global tracker which merges all tracklets belonging to an object in the whole video. Particularly, we compute online descriptor reliability over time based on their discrimination. Based on the computed discriminative descriptor weights, the global matching score over descriptors of 2 tracklets is given. Then, we apply Hungarian algorithm to optimize

Table 6. Tracking performance. The best values are printed in bold, the second best values are printed in italic.

Dataset	Method	MOTA	MOTP	GT	MT	PT	ML	FG
PETS2015	Chau <i>et Al.</i> [53]	–	–	2	0.0	100.0	0.0	2
	Ours (Proposed approach + [53])	–	–	2	100.0	0.0	0.0	1
PETS2009	Chau <i>et Al.</i> [53]	0.62	0.63	21	–	–	–	8
	Bae <i>et Al.</i> with all [48]	0.83	0.69	23	100	0	0.0	4
	Zamir et Al. [95]	0.90	0.69	21	–	–	–	–
	Bae et Al.-global association [48]	0.73	0.69	23	100	0	0.0	12
	Badie et Al. [47]	0.90	0.74	21	–	–	–	–
	Badie et Al. [47] + [53]	0.85	0.71	21	66.6	23.9	9.5	6
	Ours (Proposed approach + [53])	0.86	0.72	21	76.2	14.3	9.5	4
TUD-Stadtmitte	Milan <i>et Al.</i> [74]	0.71	0.65	9	70.0	20.0	0.0	–
	Yan <i>et Al.</i> [94]	–	–	10	70.0	30.0	0.0	–
	Chau <i>et Al.</i> [53]	0.45	0.62	10	60.0	40.0	0.0	13
	Ours (Proposed approach + [53])	0.47	0.65	10	70.0	30.0	0.0	7
TUD-Crossing	Tang <i>et Al.</i> [84]	–	–	–	53.8	38.4	7.8	–
	Chau <i>et Al.</i> [53]	0.69	0.65	11	46.2	53.8	0.0	14
	Ours (Proposed approach + [53])	0.72	0.67	11	53.8	46.2	0.0	8

tracklet matching. On the other hand, a motion model is also combined with appearance descriptors in a flexible way to improve the tracking quality. Figure 12 shows the visual explanation. In frame 137, two objects have similar appearance but move with different direction. In this case, motion descriptor is more reliable. Inversely, in frame 553, two objects go consistently together but their coat and hair's colors are different. Therefore, the appearance descriptors are more reliable than motion one.

The proposed approach gets results of tracker in [53] as input and is tested on challenge datasets. The comparable results of this tracker with other trackers from the state of the art are shown in Table 6. This paper is accepted in PETs workshop [41].

7.8. Optimizing People Tracking for a Video-camera Network

Participants: Julien Badie, François Brémond.

Keywords: tracking quality estimation, error recovering, tracklet matching

This work addresses the problem of improving tracking quality during runtime. Most state-of-the-art tracking or high-level algorithms such as event recognition have difficulties to handle erroneous inputs. This framework detects and repairs detection or tracking errors. It works in an online situation and even in the case where prior knowledge of the scene (such as contextual information or training data) is not available.

The Global Tracker (figure 13) uses tracking results (tracklets) as input and produces corrected tracklets as output.

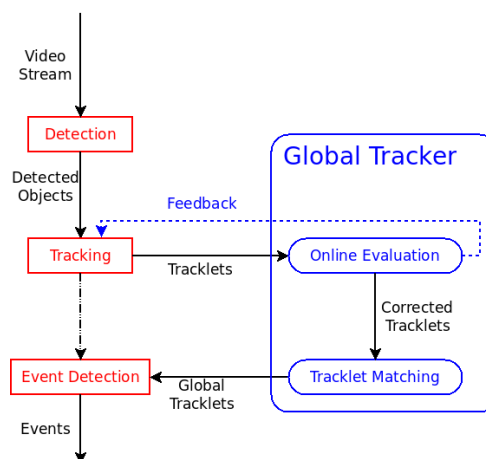


Figure 13. The Global Tracker framework, combining online evaluation and tracklet matching to improve tracking results.

The Global Tracker framework is divided into two main modules:

- Online evaluation of tracking results:** the quality of the tracking results is computed by analyzing the variation of each tracklet feature. A significant feature variation is considered as a potential error, an anomaly. To determine if this anomaly is a real error or a natural phenomenon, we use information given by the object neighborhood and the context. Finally the errors are corrected either by removing the erroneous nodes (basic approach) or by sending a signal to the tracking algorithm in order to tune its parameters for the next frames (feedback approach).

- **Tracklets matching over time:** tracklets representing the same object are merged together in a four-step algorithm. First we select key frames (frames that are close to the mean value of the features) for each tracklet. Then a visual signature is computed based on these key frames. The distance between each pair of signature is then computed. Finally the tracklet merging is done using unsupervised learning and a constrained clustering algorithm where all tracklets representing the same object are put in the same cluster.

This approach has been tested on several datasets such as PETS2009 (table 7), CAVIAR (table 8), TUD, I-LIDS and VANAHEIM and with different kinds of scenarios (tracking associated with a controller, 3D camera, camera network with overlapping or distant cameras). In each case, we are able to reach or outperform the state-of-the-art results.

Table 7. Tracking results on sequence S2.L1.View1 of the PETS2009 dataset

Methods	MOTA	MOTP
Zamir et al. [95]	0.90	0.69
Milan et al. [75]	0.90	0.74
Online evaluation	0.90	0.74
Tracklet matching	0.83	0.68
Global Tracker	0.92	0.76

Table 8. Tracking results on the Caviar dataset

Methods	MT (%)	PT (%)	ML (%)
Li et al. [73]	84.6	14.0	1.4
Kuo et al. [70]	84.6	14.7	0.7
Online Evaluation	82.6	11.7	5.7
Tracklet matching	84.6	9.5	5.9
Global Tracker	86.4	8.3	5.3

This approach is described more in detail in the PhD manuscript [27].

7.9. Multi-camera Multi-object Tracking and Trajectory Fusion

Participants: Kanishka Nithin Dhandapani, Thi Lan Anh Nguyen, Julien Badie, François Brémond.

Keywords: Multicamera, Tracklet association, Trajectory fusion, Object Tracking.

In spite of number of solutions that exist for multi-object tracking, it is still considered most challenging and unsolved computer vision problems, mainly due to inter and intra-occlusions, inferior visibility in crowded scenes, object re-entry, abrupt movement of object, placement of cameras and other detection inaccuracies that occur in single camera. Such drawbacks in single-camera multi-target tracking can be solved to an extent by obtaining more visual information on the same scene (more cameras). Few works done in the past years are [50], [79], [65], [59]. However they have their own problems such as not real run time performance, complex optimizations, hypothesizing 3D reconstruction and data association together might lead to suboptimal solutions.

We present a multi-camera tracking approach that associates and performs late fusion of trajectories in a centralized manner from distributed cameras. We use multiple views of the same scene to recover information that might be missing in a particular view. For detection we use background subtraction followed by discriminatively Trained Part Based Models . For object tracking, we use an object appearance-based tracking algorithm introduced by Chau et al [54] that combines a large set of appearance features such as 2D size, 3D displacement, colour histogram, and dominant colour to increase the robustness of the tracker to manage occlusion cases. Each camera in the network runs the detection and tracking chain independent of each

other in a distributed manner. After a batch of frames, the data from each camera is gathered to a central node by projecting the trajectories of people to the camera with the most inclusive view through a planar homography technique and then global association and fusion are performed. Unlike the temporally local (frame to frame) data-association method, global data association has ability to deal with challenges posed by noisy detections. Global association also increases the temporal stride under optimization, therefore more stable and discriminative properties of targets can be used. Trajectory similarities are calculated as heuristically weighted combination of individual features based on geometry, appearance and motion. Association is modeled as a complete K-partite graph (all pairwise relationships inside the temporal window are taken into account) K corresponds to number of cameras in network. For simplicity purpose, we use K=2. Since we use complete K-partite graph, we have an optimal solution. Whereas methods that model association as complex multivariate optimization, upon scaling, face the problem of being stuck at local minima and may provide sub-optimal solutions. Fusion is performed using adaptive weighting method. Where the weights are derived from reliability attribute of each tracker. This enables correct and consistent trajectories after fusion even if the individual trajectories have inherent noises, occlusion and false positives

Our approach is evaluated on the publicly available PETS2009 dataset. PETS2009 is a challenging dataset due to its low FPS and interobject occlusions . We choose View1,View3 and View5 in S2.L1 scenario to evaluate. The results can be seen in Figure 14

The results are encouraging and are very raw and preliminary with lot of scope and room for improvement. With more fine tuning, error rate can be improved. However too significant errors in people detection to build on top of it. Thus, we need training detector on specific datasets to improve the approach. As future work, we will study if we can improve the optimization stage with a more complex optimization using minimal graph flow would improve the results drastically.

Scenario	Method	Camera ID's	MOTA(%)	MOTP(%)	MT(%)	PT(%)	ML(%)
PETS 2009 S2.L1	Berclaz et al.[1]	1,3,5,6,8	82	56	-	-	-
	Leal-Taixe et al.[2]	1,5	76	60	-	-	-
	Leal-Taixe et al.[2]	1,5,6	71.4	53.4	-	-	-
	Murray Evans et al[4]	2 Cameras	63	55			
	Martin Hofmann et al[3]	1,5	99.4	82.9	100	0	0
	Martin Hofmann et al.[3]	1,5,7	99.4	83.0	100	0	0

Scenario	Method	Camera ID's	MOTA(%)	MOTP(%)	MT(%)	PT(%)	ML(%)
PETS2009 S2.L1	Our Approach	1,3	76.33	65.28	92.59	0.035	0.714

Figure 14. Result of our approach on PET2009 dataset.

7.10. Person Re-identification in Real-World Surveillance Systems

Participants: Furqan Mohammad Khan, François Brémond.

Keywords: re-identification, long term visual tracking, signature modeling

Cost of supervised metric learning Person re-identification problem has recently received a lot of attention and the recent focus is to use supervised model training to learn cross camera appearance transformation. In general, $O(n^2)$ models are trained in a surveillance network with n cameras, one for each camera pair. $2p$ tracks are required to train one model with p person identities. In a real-world surveillance network with non-overlapping fields of view, a person appears only in a subset of cameras (see figure 15 , courtesy of [51]). This puts the requirement of number of tracks to train all models at $O(n^2p)$, or more precisely, $n(n-1)p$. That

is, to train each model with 100 people in a 10 camera network we need 9000 tracks. For supervised training, these tracks need to be given consistent identities, and worse, have their bounding boxes marked. This is a significant burden on human annotators for deployment in real-world. Further, the annotation cost has to be repaid at a significant fraction if only one new camera is added to the system (may be due to failure of an existing camera), or if the lighting changes significantly (in case of outdoor surveillance). In our opinion, this is a significant bottleneck for supervised metric learning based re-identification in real-world.

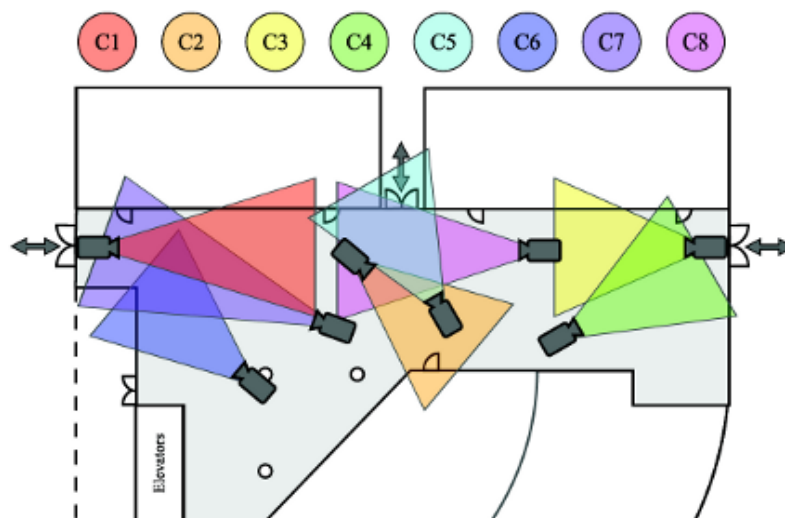


Figure 15. Camera arrangement in multi-camera surveillance scenario of SAIVT-SoftBio dataset [51]

Improved re-identification through signature modeling Re-identification is challenging because variance is intra-class appearance is often higher than inter-class appearance due to varying lighting conditions and viewpoints, and non-uniqueness of clothing. More importantly, in real-world when re-identification is fed by automated human detectors and trackers, significant mis-alignment or partial visibility of the person within proposed bounding box make it difficult to extract relevant features. Our work focuses on improving signature construction from low level features for multi-shot re-identification. We explicitly model multi-modality of person appearance using a feature mixture (corresponding publication is under review at this moment). This improves state-of-the-art re-identification performance on SAIVT-SoftBio [51] dataset and performs equally well as state-of-the-art metric learning methods on iLIDS-VID [88] and PRID2011 [64] datasets. The performance comparison of our method with state-of-the-art is presented using CMC in figure 16 (our results are denoted by MCAM).

7.11. Human Action Recognition in Videos

Participants: Piotr Bilinski, François Brémond.

Keywords: Action Recognition, Video Covariance Matrix Logarithm, VCML, Descriptor
Video Covariance Matrix Logarithm for Human Action Recognition in Videos

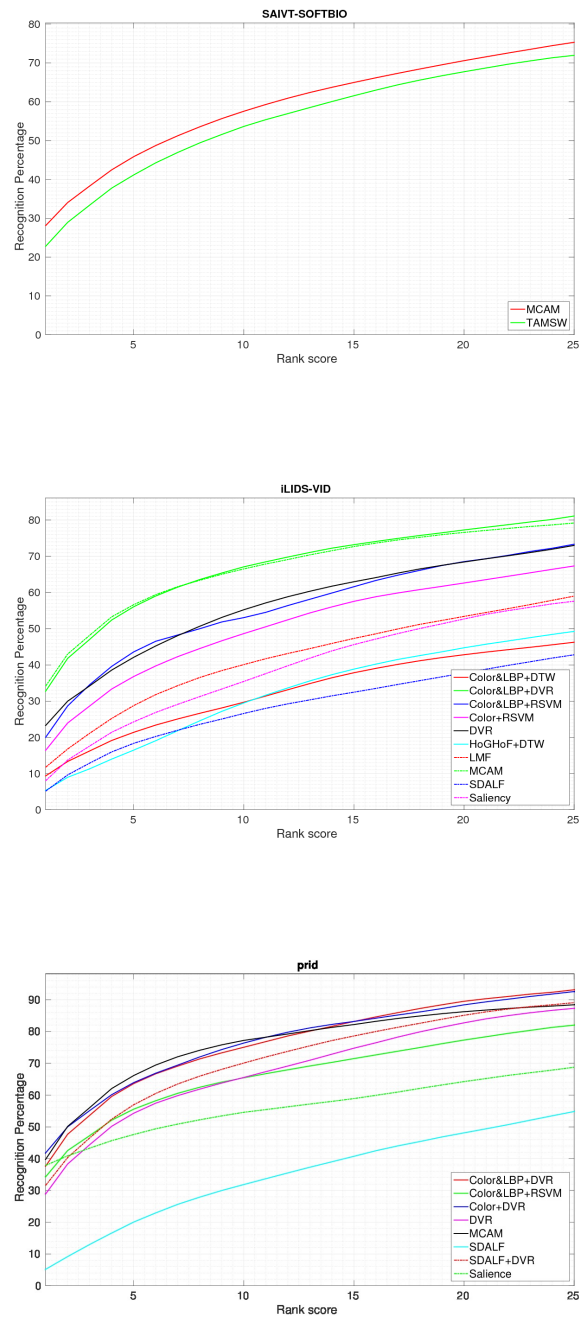


Figure 16. Performance comparison of our MCAM approach using CMC curves on different datasets. **Top:** Comparison with TAMSW [49] on SAIVT-SoftBio dataset; **middle:** Comparison with Color+RSVM [88], Color&LBP+DTW [88], Color&LBP+DVR [88], Color&LBP+RSVM [88], DVR [88], HoG-HoF+DTW [88], LMF [97], Saliency [96], and SDALF [60] on iLIDS-VID dataset; **bottom:** Comparison with Color+DVR [88], Color&LBP+DVR [88], Color&LBP+RSVM [88], DVR [88], Saliency [96], and SDALF [60] on PRID2011 dataset.

In this work, we propose a new local spatio-temporal descriptor for videos and we propose a new approach for action recognition in videos based on the introduced descriptor. Overview of the proposed action recognition approach based on the introduced descriptor is presented in Figure 17. The new descriptor is called the Video Covariance Matrix Logarithm (VCML). The VCML descriptor is based on a covariance matrix representation, and it models relationships between different low-level features, such as intensity and gradient. We apply the VCML descriptor to encode appearance information of local spatio-temporal video volumes, which are extracted by the (Improved) Dense Trajectories. Then, we present an extensive evaluation of the proposed VCML descriptor with the (Improved) Fisher Vector encoding and the Support Vector Machines on four challenging action recognition datasets (*i.e.* URADL, MSR Daily Activity 3D, UCF50, and HMDB51 datasets). We show that the VCML descriptor achieves better results than the state-of-the-art appearance descriptors. In comparison with the most popular visual appearance descriptor, *i.e.* the HOG descriptor, the VCML achieves superior results. Moreover, we present that the VCML descriptor carries complementary information to the HOG descriptor and their fusion gives a significant improvement in action recognition accuracy (*e.g.* the VCML improves the HOG by 15% on the HMDB51 dataset). Finally, we show that the VCML descriptor improves action recognition accuracy in comparison to the state-of-the-art (Improved) Dense Trajectories, and that the proposed approach achieves superior performance to the state-of-the-art methods. The proposed VCML based technique achieves 94.7% accuracy on the URADL dataset, 85.9% on the MSR Daily Activity 3D dataset, 92.1% on the UCF50 dataset, and 58.6% on the HMDB51 dataset. More results and comparisons with the state-of-the-art are presented in Table 9 and Table 10. To the best of our knowledge, this is the first time covariance based features are used to represent the trajectories. Moreover, this is the first time they encode the structural information and they are applied with the (Improved) Fisher Vector encoding for human action recognition in videos. This work has been published in [40].

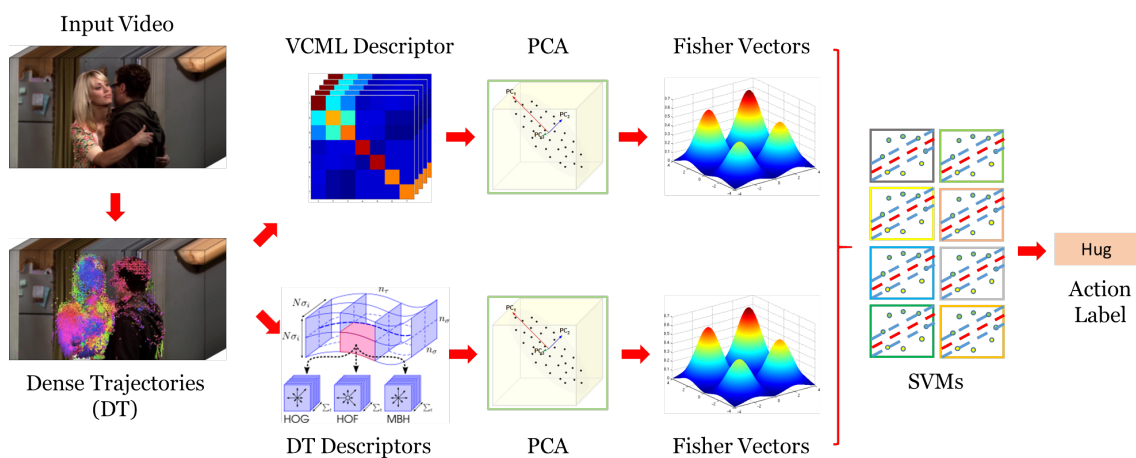


Figure 17. Overview of our action recognition approach based on the introduced VCML descriptor.

7.12. Evaluation of Event Recognition without Using Ground Truth

Participants: Ramiro Diaz, Carlos Fernando Crispim Junior, François Brémond.

Keywords: Computer Vision, Event Recognition, Video Summarization.

The main goal of the work is to improve the Event Recognition process and to improve the way we build the event models as well. The work concerns the Valrose Nursing Home, it consists in monitoring older people with health issues like Dementia and who are in need of care and stimulation.

Table 9. Comparison with the state-of-the-art on the URADL and MSR Daily Activity 3D datasets. The table presents the accuracy of our approach using Dense Trajectories (DT) and Improved Dense Trajectories (IDT).

URADL		MSR Daily Activity 3D	
Benabbas <i>et Al.</i> , 2010	81.0	Koperski <i>et Al.</i> , 2014	72.0
Raptis and Soatto, 2010	82.7	JPF – Wang <i>et Al.</i> , 2012	78.0
Messing <i>et Al.</i> , 2009	89.0	Oreifej and Liu, 2013	80.0
Bilinski and Bremond, 2012	93.3	AE – Wang <i>et Al.</i> , 2012	85.7
Dense Trajectories	94.0	Dense Trajectories	76.2
Our Approach (DT)	94.0	Our Approach (DT)	78.1
Our Approach (IDT)	94.7	Our Approach (IDT)	85.9

Table 10. Comparison with the state-of-the-art on the UCF50 and HMDB51 datasets. The table presents the accuracy of our approach using Dense Trajectories (DT) and Improved Dense Trajectories (IDT).

UCF50		HMDB51	
Kantorov and Laptev, 2014	82.2	Kantorov and Laptev, 2014	46.7
Shi <i>et Al.</i> , 2013	83.3	Jain <i>et Al.</i> , 2013	52.1
Oneata <i>et Al.</i> , 2013	90.0	Oneata <i>et Al.</i> , 2013	54.8
Wang and Schmid, 2013	91.2	Wang and Schmid, 2013	57.2
Dense Trajectories	84.2	Dense Trajectories	47.0
Our Approach (DT)	88.1	Our Approach (DT)	52.9
Our Approach (IDT)	92.1	Our Approach (IDT)	58.6

Since the video dataset contains data for about 8 months, a new evaluation method is required to properly analyze the whole dataset and gain a deeper understanding of it. Our approach consists in displaying the data in a way that can be useful either for doctors, as well for engineers to detect failures and to improve the event recognition process in an efficient way. Because of this need, a new evaluation tool has been developed and named Event Plotter.

This tool provides a new method for event evaluation. First of all, as we do not have ground truth information for the total duration of the 8 months, but just for one week, so another method is needed to check the event model efficiency. To address this issue, the tool displays all the events in the desired time period (as clusters on a timeline) and single events (or time intervals) can be selected to quickly check the video and visualize the results of the event recognition working on the fly –see Figure 18 -. The goal of this work is to understand how event recognition works, change the models on the fly, import them, and see how the recognition changes in real time.

Also, to compare new event models with the old ones, video summarization is implemented as well. Event based video summarization is utilized here to check how the recognition of one particular event type changes globally on the whole video and to display the recognition results. Also video summarization can be useful for doctors to check the way patients behave, for example playing all the videos of event "Get-up from bed", trying to predict patterns.

The data processed to address this issue was 1 week, because it was the time corresponding to the ground truth data. With these processed data, we have tested the efficiency of the Event Plotter tool, and we are currently improving the event recognition process by changing event models, adding new zones, and testing them on the fly.

7.13. Semantic Event Fusion of Different Visual Modality Concepts for Activity Recognition

Participants: Carlos Fernando Crispim-Junior, François Brémond.

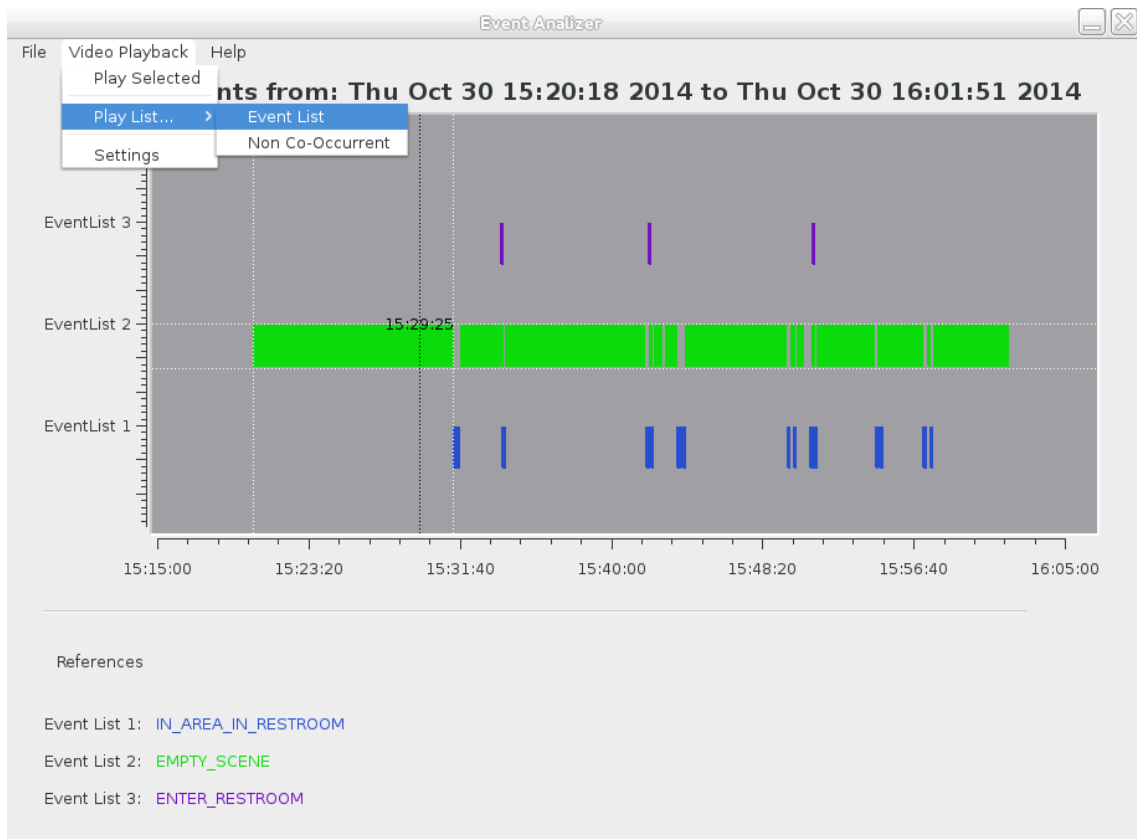


Figure 18. GUI of Event Plotter with 3 loaded Event Lists.

Keywords: Knowledge representation formalism and methods, Uncertainty and probabilistic reasoning, Concept synchronization, Activity recognition, Vision and scene understanding, Multimedia Perceptual System,

Combining multimodal concept streams from heterogeneous sensors is a problem superficially explored for activity recognition. Most studies explore simple sensors in nearly perfect conditions, where temporal synchronization is guaranteed. Sophisticated fusion schemes adopt problem-specific graphical representations of events that are generally deeply linked with their training data and focus on a single sensor. In this work we have proposed a hybrid framework between knowledge-driven and probabilistic-driven methods for event representation and recognition. It separates semantic modeling from raw sensor data by using an intermediate semantic representation, namely concepts. It introduces an algorithm for sensor alignment that uses concept similarity as a surrogate for the inaccurate temporal information of real life scenarios (Fig. 20). Finally, it proposes the combined use of an ontology language, to overcome the rigidity of previous approaches at model definition, and a probabilistic interpretation for ontological models, which equips the framework with a mechanism to handle noisy and ambiguous concept observations, an ability that most knowledge-driven methods lack (Fig. 19). We evaluate our contributions in multimodal recordings of elderly people carrying out instrumental activities of daily living (Table 11). Results demonstrated that the proposed framework outperforms baseline methods both in event recognition performance and in delimiting the temporal boundaries of event instances

This work has been developed as a collaboration between different teams in Dem@care consortium (Inria, University of Bordeaux, and CERTH). We thank the other co-authors for their contributions and support in the development of this work up to its submission for publication.

Table 11. Comparison to baseline methods in the test set

mean F_1 -score	Fusion approach		
	Baselines		Ours
IADL	SVM	OSF	
S. bus line	44.19	31.36	73.10
M.finances	43.99	0.00	43.73
P. pill box	45.86	49.11	65.02
P. drink	20.02	24.29	64.03
Read	90.18	91.82	95.22
T.telephone	72.12	0.00	75.58
W. TV	2.32	0.00	35.80
W. Plant	0.00	0.00	100.00
Average	39.83	24.57	69.06

OSF: Ontology-based Semantic Fusion

7.14. Semi-supervised Activity Recognition Using High-order Temporal-composite Patterns of Visual Concepts

Participants: Carlos Fernando Crispim-Junior, Michal Koperski, Serhan Cosar, François Brémond.

Keywords: visual concepts, semi-supervised activity recognition, complex activities, cooking composite activities

Methods for action recognition have evolved considerably over the past years and can now automatically learn and recognize short term actions with satisfactory accuracy. Nonetheless, the recognition of activities - a composition of actions and scene objects - is still an open problem due to the complex temporal, composite structure of this category of events. Existing methods either still focus on simple activities or oversimplify the modeling of complex activities by only targeting whole-part relations between activity components, like actions. In this work, we have investigated a hierarchical, semi-supervised approach that unsupervisedly learns actions from the composite patterns of atomic concepts (*e.g.*, slice, tomato), and complex activities from the

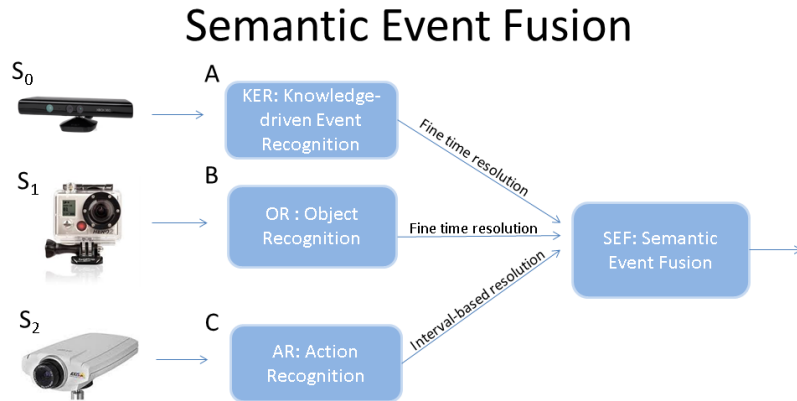


Figure 19. Semantic event fusion framework: detector modules (A-C) process data from their respective sensors (S_0 - S_2) and output concepts (objects and low-level events). Semantic Event Fusion uses the ontological representation to initialize concepts to event models and then infer complex, composite activities. Concept fusion is performed on millisecond temporal resolution to cope with instantaneous errors of concept recognition.

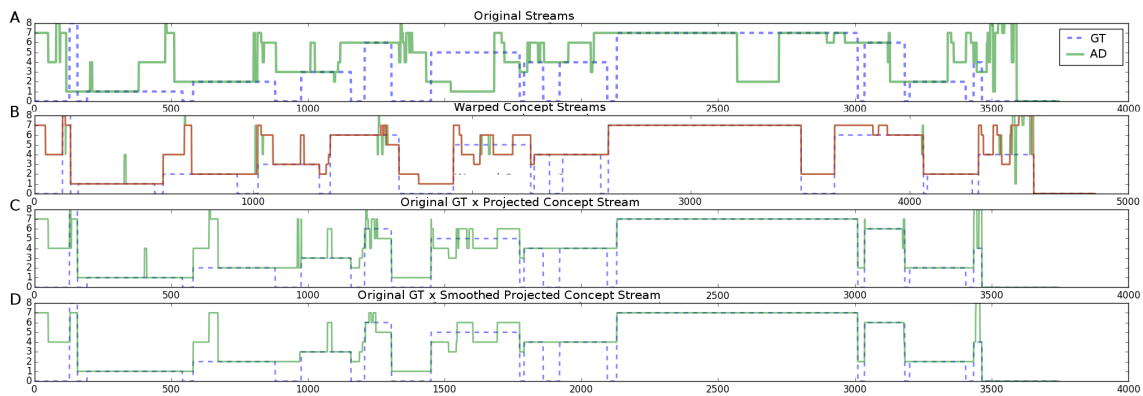


Figure 20. Semantic alignment between the concept stream of the action recognition detector (AR) and a concept stream (GT) generated from events manually annotated by domain experts using the time axis of the color-depth camera. X-axis denotes time in frames, and Y-axis denotes activity code (1-8), respectively, search bus line on the map, establish bank account balance, prepare pill box, prepare a drink, read, talk on the telephone, watch TV, and water the plant. From top to bottom, images denote: (A) original GT and AR streams, (B) GT and AR streams warped, AR stream warped and smoothed (in red), (C) original GT and AR stream warped and then backprojected onto GT temporal axis, (D) original GT and AR warped, backprojected, and then smoothed with median filtering.

temporal patterns of concept compositions (actions). On a first step, an unsupervised, inductive approach iteratively builds a multi-scale, temporal-composite model of the concept occurrences during the activity taking place (Fig. 22). Then, activity recognition is performed by comparing the similarity of the generated model of a given video and *a priori* learned and labeled unsupervised models. We have evaluated the proposed method in the MPII Cooking Composite Activities dataset (Fig. 21), a video collection where people perform a set of complex activities related to cooking recipes. To tackle this dataset it is necessary to recognize a large variety of visual concepts (*e.g.*, from actions, such as cutting and stirring, to objects, such as tomato and cutting board). Moreover, the detection of cooking activities is a very challenging problem since we observe a low inter-class variance between activity classes, and a high intra-class variance within an activity due to person to person differences in performing them. The proposed approach presents a mean average precision (mAP) of $56.36\% \pm 5.1\%$, and then outperforms previous methods ([81], mean AP 53.90%). This improvement is devoted to the modeling of deeper composite and temporal relations between visual concepts (from 2nd to 5th order compositions). The performance of the proposed method is mostly limited by the performance of low-level concept detectors. Future work will investigate ways to extend the current probabilistic model to handle more efficiently the differences in concept detector performance.



Figure 21. Illustration of one of the cooking recipes of Cooking Composite dataset [81]

7.15. From Activity Recognition to the Assessment of Seniors' Autonomy

Participants: Carlos Fernando Crispim-Junior, Carola Strumia, Alvaro Gomez Uria Covella, Alexandra Konig, François Brémond.

Activity recognition plays a fundamental role in several research fields as a way to extract semantic meaning from images and videos, to find more accurate matches for textual queries in video search engines, and to analyze long-term activity patterns in assisted living scenarios, such as seniors living at home. In this sense, we have continued our work on activity monitoring by proposing a novel knowledge-based event monitoring system that combines the observations of a vision system with expert knowledge and scene semantics, to recognize daily living activities in assisted living scenarios.

The approach's novelty lies in the combination of a flexible constraint-based ontology language for event modeling with efficient and robust algorithms to detect, track and re-identify people using color-depth sensing (low-level vision). The robust low-level vision promotes the modeling of longer and more complex events,

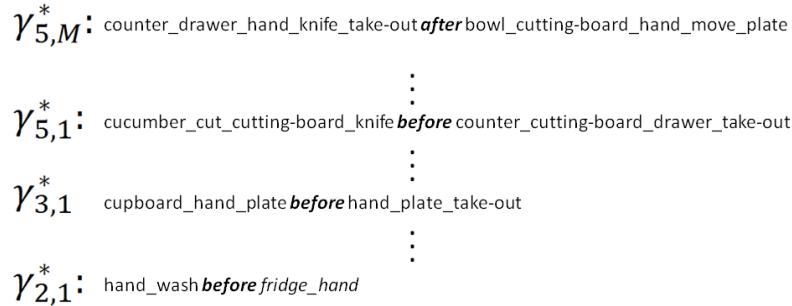


Figure 22. Example of temporal concept composite patterns of different arities extracted from an unlabeled activity.

while the ontology language provides a flexible way to describe event and incorporate domain knowledge, and ease knowledge transfer across different scenes. The proposed approach has been investigated for two assisted living scenarios: a) the monitoring of physical tasks and daily living activities in observation rooms of hospital and clinics, and b) daily and nightly activities of seniors living in nursing home apartments. To evaluate our approach performance compared to state of art methods, we have computed its results for GAADR dataset. This is public dataset, which is composed of videos of seniors performing physical tasks and activities of daily living. Evaluation results (Table 12) have demonstrated that our approach achieves an average F_1 - score 20 % higher than the baseline method [89].

Table 12. Recognition of IADLs - GAADR data set - F 1 -score

Event	DT-HOG	DT-HOF	DT-MBHy	Proposed
Account Balance	44.96	34.71	42.98	66.67
Prepare Drink	81.66	44.87	52.00	100.00
Prepare Drug Box	14.19	0.00	0.00	57.14
Read Article	52.10	42.86	33.91	63.64
Talk on telephone	82.35	0.00	33.76	100.00
Turn on radio	85.71	42.52	58.16	94.74
Water Plant	0.00	0.00	0.00	52.63
Average \pm SD	51.8 \pm 34.4	23.6 \pm 22.3	31.5 \pm 23.3	76.4 \pm 21.0

Given the satisfactory performance of the proposed activity recognition framework we have also investigated it as a method to automatically measure a seniors' autonomy in quantitative and objective fashion. To do so, we have developed a probabilistic model that takes as input the recognized activities and gait-patterns from the period of time the person performs physical tasks. The proposed autonomy model has presented an average performance of 83.67 %, which suggests that the use of such technologies may provide clinicians with diagnostic relevant information, and decrease observer's biases when compared to clinical scales. The results of this investigation have been published in [33].

7.16. Serious Games Interfaces Using an RGB-D Camera : Results and Perspectives

Participants: Baptiste Fosty, François Brémond.

Keywords: RGB-D camera analysis, walking speed, serious games, startup project

Within the context of the development of serious games for people suffering from Alzheimer disease (Az@Game project), we have developed algorithms to interact with the virtual environment through simple gesture recognition and walking speed computation. We have shown in previous work that the walking speed measured by our system is accurate enough within this context and reproducible. A paper has been submitted in Gait and Posture journal (now in reviewing process).

Concerning the gesture recognition algorithm, it consists in recognizing three basic gestures (right arm left on the side, left arm left on the side, right or left arm left on top). We performed a small experimentation to test the robustness of the system in detecting these gestures where participants (10 in total) had to perform 10 times each gesture while walking at 2.5km/h on the treadmill (see Figure 23). The results are shown in Table 13 .

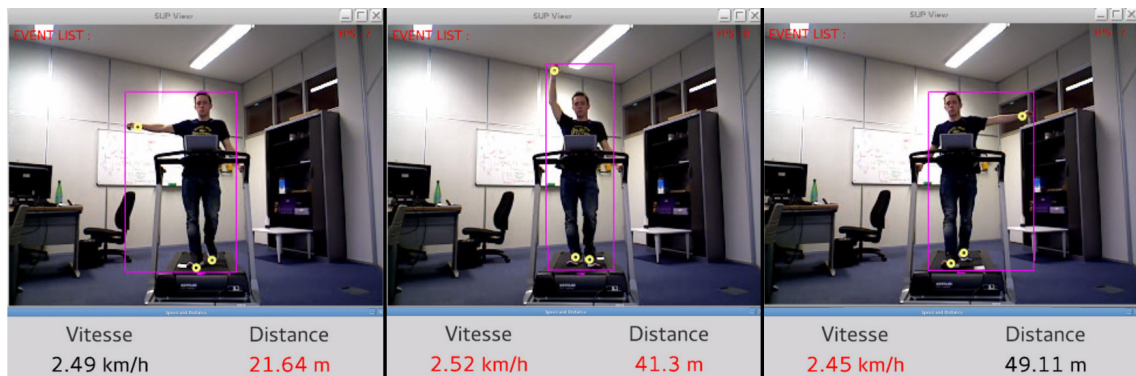


Figure 23. Display of gesture recognition and speed computation. Distance is red when right arm is left, speed is red when left arm is left, both are red when one arm is on top. Yellow dots correspond to feet and arm detection.

Table 13. Gesture recognition results.

	Right	Left	Top	TOTAL
Recall	97.0	95.0	100.0	97.3
Precision	100.0	97.9	94.3	97.3
F-Score	98.5	96.4	97.1	97.3

Following that, we decided to study whether this system would be useful in rehabilitation. Some experts of this domain in a rehabilitation center (Centre H elio Marin in Vallauris, France) have been interviewed and they were very enthusiasts about using this type of system to get objective gait parameters. To go further on the market opportunity and evaluate the feasibility of a technology transfer, we studied the concurrent products and contacted more than thirty other rehabilitation centers to have a deeper understanding of the needs and validate our idea. This investigation lead to the proposition of a startup project(BOMOTECH) to Inria which has been accepted and funded for the next 7 months, during what the goal is to get closer from a marketable tool.

7.17. Assistance for Older Adults in Serious Game Using an Interactive System

Participants: Minh Khue Phan Tran, Fran ois Br emond, Philippe Robert.

Keywords: Older Adults, Assistance, Serious Games

Serious Games offer a new way to older adults to improve their abilities such as vision, balance or memory. However, cognitive impairment causes a lot of difficulties to them when actively practising these games. Their engagement and motivation are reduced rapidly when encountering successive problems without any help. Our hypothesis is that this problematic situation can be handled if they are assisted regularly. We propose then an interactive system which can determine dynamically the situations and provide different helps in real-time. We focus on two main problems that the older players encounter regularly :

- they forget how to continue to play the game.
- they make a lot of errors.

The system determines the above problems by computing various characteristics of the player (skeleton, postures, gestures ...) along the game states. This process is presented in Figure 24 . The characteristics of the player, which are collected by the Recognition Module thanks to Kinect Camera and the related SDK, are sent to the Interaction Module. This module associates these data with the game states provided by the game in order to recognize the problem and interacts with the player through a 3D-animated avatar.

The system is tested with 3 groups of patients described by 3 different cognitive states : mnesic plaits, MCI and Alzheimer. The patients are invited to play a concentration-based game with a Kinect camera. Each patient plays 3 phases : playing with therapist, playing alone and playing with the avatar. The playing time and the final score of each game phase are recorded. Here, the system takes into account the player's gestures and the game states for recognizing two situations :

- the player reacts too late and too slowly to the current game task.
- the player makes many mistakes.

The experimental results confirmed our hypothesis. Most of the patients have the best performance in phase "playing with the avatar". Their playing time is shorter and their final score is higher in phase "playing with assistance" than in phase "playing alone". The results are presented in the publication [36] accepted by the Games and Learning Alliance Conference in December 2015. Future work aims at improving the system and compare its efficiency with the one of "humans assistances".

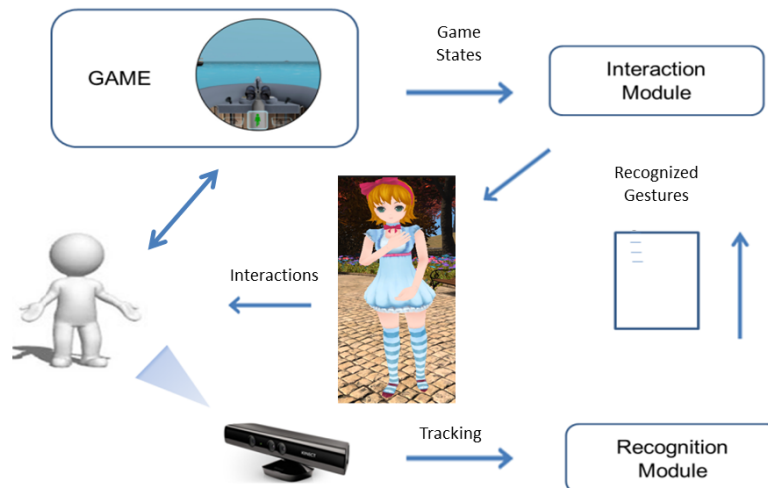


Figure 24. Assisting older adults in serious game playing

7.18. Generating Unsupervised Models for Online Long-Term Daily Living Activity Recognition

Participants: Farhood Negin, Serhan Coşar, Michal Koperski, François Brémond.

Keywords: Unsupervised Activity Recognition

Generating Unsupervised Models for Online Long-Term Daily Living Activity Recognition

In this work, we propose an unsupervised approach that offers a comprehensive representation of activities by modeling both global and body motion of people. Compared to existing supervised approaches, our approach automatically learns and recognizes activities in videos without user interaction. First, the system learns important regions in the scene by clustering trajectory points. Then, a sequence of primitive events is constructed by checking whether people are inside a region or moving between regions. This enables to represent the global movement of people and automatically split the video into clips. After that, using action descriptors [90], we represent the actions occurring inside each region. Combining action descriptors with global motion statistics of primitive events, such as time duration, an activity model that represents both global and local action information is constructed. Since the video is automatically clipped, our approach performs online recognition of activities. The contributions of this work are twofolds: (i) generating unsupervised human activity models that obtains a comprehensive representation by combining global and body motion information, (ii) recognizing activities online and without requiring user interaction. Experimental results show that our approach increases the level of accuracy compared to existing approaches. Figure 25 illustrates the flow of the system.

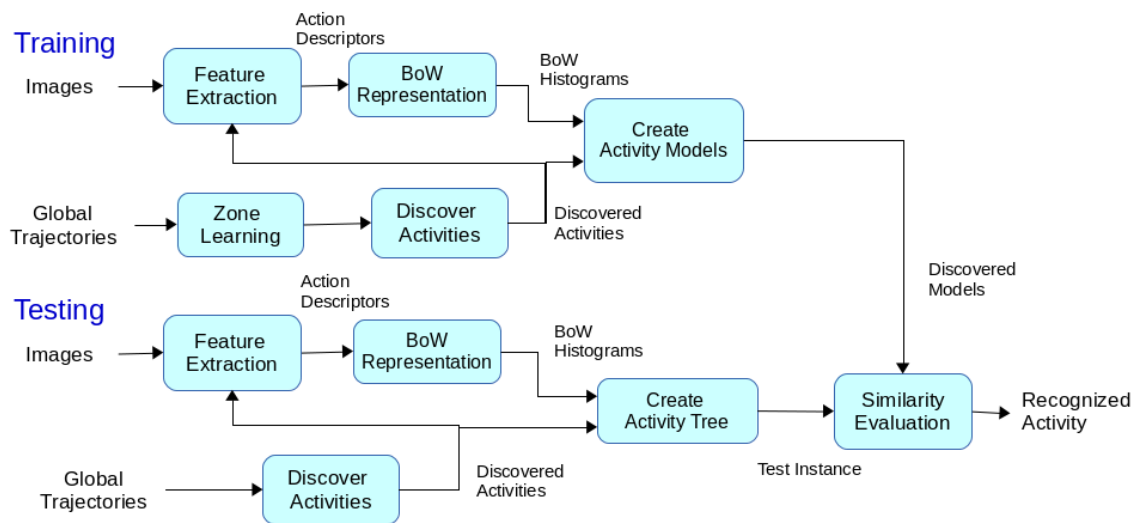


Figure 25. Architecture of the framework: Training and Testing phases

The performance of the proposed approach has been tested on the public GAADR dataset [67] and CHU dataset (<http://www.demcare.eu/results/datasets>) that are recorded under EU FP7 Dem@Care Project1 in a clinic in Thessaloniki, Greece and in Nice, France, respectively. The datasets contain people performing everyday activities in a hospital room. The activities considered in the datasets are listed in Table 1 and Table 2. Each person is recorded using RGBD camera of 640x480 pixels of resolution. The GAADR dataset contains 25 videos and the CHU dataset contains 27 videos. Each video lasts approximately 10-15 minutes.

We have compared our approach with the results of the supervised approach in [90]. We did also a comparison with an online supervised approach that follows [90]. For doing this, we train the classifier on clipped videos and perform the testing using sliding window. In the online approach, a SVM is trained using the action descriptors extracted from groundtruth intervals. We have also tested different versions of our approach that i) only uses global motion features and ii) which only uses body motion features. We have randomly selected 3/5 of the videos in both datasets for learning the activity models. The codebook size is set to 4000 visual words for all the methods.

The performance of the online supervised approach and our approach in GAARDR dataset are presented in Table 1. In all approaches that use body motion features, HoG descriptors are selected since they give the best results. It can be clearly seen that, using models that represent both global and body motion features, our unsupervised approach enables to obtain high sensitivity and precision rates. Compared to the online version of [90], thanks to the learned zones from positions and discovered activities, we obtain better activity localization, thereby better precision. However, since the online version of [90] utilizes only dense trajectories (not global motion), it fails to localize activities. Hence, it detects the intervals that does not include an activity (e.g. walking from radio desk to phone desk) and for "prepare drug box", "watering plant", and "reading" activities, it cannot detect the correct intervals of the activities. Compared to the unsupervised approach that either use global motion features or body motion features, we can see that, by combining both features, our approach achieves more discriminative and precise models, thereby improves both sensitivity and precision rates. By combining global and body motion features, our approach benefits from discriminative properties of both feature types. Table 1 also presents the results of the supervised approach in [90]. Although the supervised approach uses groundtruth intervals in test videos in an offline recognition scheme, it fails to achieve accurate recognition. As our approach learns the zones of activities, we discover the places where the activities occur, thereby we achieve precise and accurate recognition results. Since this information is missing in the supervised approach, it detects "turning on radio" while the person is inside drink zone preparing drink.

Table 2 shows the results of the online supervised approach and our approach in CHU dataset. MBH descriptor along y axis and HoG descriptor gives the best results for our approach and the online supervised approach, respectively. In this dataset, since people tend to perform activities in different places (e.g. preparing drink at phone desk), it is not easy to obtain high precision rates. However, compared to the online version of [90], our approach detects all activities and achieves a much better precision rate. The online version of [90] again fails to detect activities accurately, thereby misses some of the "preparing drink" and "reading" activities and gives many false positives for all activities.

Thanks to the activity models learned in unsupervised way, we accurately perform online recognition. In addition, the zones learned in an unsupervised way help to model activities accurately, thereby most of the times our approach achieves more accurate recognition compared to supervised approaches. This paper has been published in third Asian Conference on Pattern Recognition (ACPR 2015) [35].

7.19. Run-time Adaptation of Video Systems

Participants: Sabine Moisan, Jean-Paul Rigault, François Brémond.

In the framework of our research on model engineering techniques for video-surveillance systems, we have focused this year on run-time reconfiguration of such systems. The goal is to follow the "model at run-time" approach and to obtain context-aware self-adaptive video systems. In this approach models are kept and used at run-time. In our case, these models describe all the possible run-time configurations. They are specified using *Feature Models*.

Run-time reconfiguration means to react to context changes by tuning, adding, removing, or replacing components of the video chain, and possibly changing the chain itself.

So far, we have defined a run-time architecture consisting of three layers. The lower level describes the video analysis components and the context events. The upper one handles feature model adaptation. The middle layer is an adapter: it transforms lower level context event occurrences into upper level feature reconfiguration; in the other direction, it transforms the corresponding feature reconfigurations into video components reconfigurations.

Table 14. The activity recognition results for GAADR dataset. Bold values represent the best sensitivity and precision results for each class.

ADLs	Supervised Approach [90]		Online Version of [90]		Unsupervised (Only Global Motion)		Unsupervised (Only Body Motion)		Proposed Approach	
	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)
Answering Phone	100	88	100	70	100	100	57	100	100	100
Establish Acc. Bal.	67	100	100	29	100	86	50	100	100	86.67
Preparing Drink	100	69	100	69	78	100	100	100	100	100
Prepare Drug Box	<u>58.33</u>	100	11	20	33.34	100	33.34	100	33.34	100
Watering Plant	<u>54.54</u>	100	0	0	44.45	57	33	100	44.45	100
Reading	100	100	88	37	100	100	38	100	100	100
Turn On Radio	60	86	<u>100</u>	75	89	89	44	100	89	100
AVERAGE	77.12	91.85	71.29	42.86	77.71	90.29	50.71	100	80.97	98.10

Table 15. The activity recognition results for CHU dataset. Bold values represent the best sensitivity and precision results for each class.

ADLs	Supervised Approach [90]		Online Version of [90]		Proposed Approach	
	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)	Sens. (%)	Prec. (%)
Answering Phone	57	78	100	86	100	65
Preparing Drink	78	73	92	43	100	58
Prepare Drug Box	100	83	100	43	100	100
Reading	35	100	92	36	100	78
Using Bus Map	90	90	100	50	100	47
AVERAGE	72.0	84.80	90.95	48.76	100	70.00

This year we focused on the upper layer. We first formalized the run-time feature reconfiguration rules. First, any reconfiguration should respect the semantics of feature models and their attached constraints. Second, the reconfiguration should satisfy the requests from the middle layer, essentially selections and deselections of features. From these two requirements, we identified three possible outcomes: successful reconfiguration, impossible reconfiguration (selection/deselection conflicts), and “undefined” reconfiguration (not enough information to get through the process). We also determined the actions to take in these cases. In particular, in the last two cases, we decided to let the component configuration unchanged.

To implement this upper layer, we first attempted to rely on an existing feature model manipulation framework, namely FAMILIAR [45]. However, this approach suffers from a number of drawbacks. First, FAMILIAR is a standalone Java program, whereas the rest of the system is written in C++, for performance reasons and library availability. Hence, using FAMILIAR implies superfluous back and forth inter-module communications and data transformations. Second, we confirmed that FAMILIAR is more a system deployment tool than a run-time reconfiguration one. In particular it cannot fulfill all the reconfiguration rules that we have formalized. Therefore, we are completing a full re-implementation of the upper layer.

The programming language homogeneity permits a more efficient integration of the three layers. In particular, it becomes easier to incorporate our extensions to feature models such as quality metrics [34].

7.20. Scenario Description Language

Participants: Sabine Moisan, Annie Ressouche, Jean-Paul Rigault, Nazli Temur, François Brémond.

Last year, we developed a scenario recognition engine based on the Synchronous Model of reactive systems. We now need a scenario description language friendly to our end users who are not computer scientists in general. In fact, Stars has already defined such a language. However, it is a declarative language based on (temporal) constraints. This is certainly not the most natural and the simplest way for end users to express their domain specific scenarios.

Consequently, we started this year a comparative study of different means to express scenarios in various domains (video understanding but also games, movies, music, criminology, military strategy...). We investigated 16 formalisms covering these domains. We defined a comparison grid based on criteria relevant for our video understanding goals. We retained 9 such criteria: application domain scope, ease of use, representation of scenario basic elements (background, scene, roles...), modularity (possibility of scenario hierarchy), time representation (absolute, logical, multi-clocks, no clocks...), expression of temporal constraints, representation of repetitive patterns, support for concurrency and parallelism, and finally formal foundations.

To complete the study, we conducted an experiment: describing a case study scenario using some of these formalisms to concretely estimate their advantages and drawbacks, especially their ease of use.

At this time, none of the studied languages fulfills completely our needs. Many languages are graphical ones. While this may appear as user friendly, scalability and automatic analysis become an issue. Some languages lack formal semantics, which is not acceptable in our case; others are merely extensions of computer languages, hence dedicated to specialists.

We plan to define our own version, which will rely on solid semantic foundations. (see section 7.21). To enforce user-friendliness, we started to collaborate with ergonomists.

7.21. Scenario Recognition

Participants: Annie Ressouche, Sabine Moisan, Jean-Paul Rigault, Ines Sarray, Daniel Gaffé.

Keywords: Synchronous Modeling, Model checking, Mealy machine, Cognitive systems.

For a long time, Stars strategy has been to favor the easy generation of activity recognition systems. These systems correspond to a succession of pattern matching and clustering algorithms, combined with adequate knowledge representation (e.g. scene topology, temporal constraints) at different abstraction levels (from raw signal to semantics). Due to the large range of application domains (surveillance, safety, health care, ...), we propose a generic approach to design activity recognition engines. Moreover, such domains require high dependability due to possible safety issues. Thus, our approach should also rely on formal methods to describe, analyze, verify, and generate effective recognition engines. We consider activity recognition engines as reactive systems that react to input events from their environment and produce output events in the form of alarms or notifications. Such engines are intrinsically real time, reactive and they evolve in discrete time. As a consequence, to recognize scenarios, we adapt the usual techniques of synchronous modeling approach to express scenario behaviors. This approach facilitates scenario validation and allows us to generate a recognizer for each scenario.

Our previous developments, on top of existing synchronous languages as Lustre and LE (see section 7.22), were convenient for rapid prototyping. However, even if LE is not a closed environment, it appeared as difficult as Lustre to customize, for efficiency reasons. This year, in the framework of Ines Sarray PhD thesis, we began to define a synchronous semantics for the future scenario language (see section 7.20). The idea is to generate automatically recognition engines at compilation time. The compilation itself is totally based on the semantics. To complete this approach we will rely on both our experiment with the LE language last year and on the LE compilation process.

7.22. The Clem Workflow

Participants: Annie Ressousche, Daniel Gaffé, Imane Khalis.

Keywords: Synchronous languages, Synchronous Modeling, Model checking, Mealy machine.

This research axis concerns the theoretical study of a synchronous language LE –with modular compilation– and the development of a toolkit (see Figure 26) around the language to design, simulate, verify, and generate code for programs. The novelty of the approach is the ability to manage both modularity and causality.

This year, we continued to focus on the improvement of both LE language and compiler concerning data handling and the generation of back-ends, required by other research axis of the team (see 7.21 and 7.23). We also improved the design of a new simulator for LE programs which integrates our new approach. In CLEM we generate an independent intermediate code (LEC) before specific target generations. This code represents the semantics of programs with 4-valued equation systems. In our design flow, we need to simulate programs at this level. Last year, we begun to develop such a simulator in order to integrate the data part of the language. The simulator GUI has been designed in Qt and the simulator takes into account the values carried by signals. This year, during her internship, Imane Khalis has completed the simulator to allow an external computation of data values and a communication with the simulator through a socket mechanism. With this last development, the LEC simulator is complete and is integrated in the CLEM toolkit.

7.23. Safe Composition in WCOMP Middleware for Internet of Things

Participants: Annie Ressousche, Daniel Gaffé, Ines Sarray, Jean-Yves Tigli.

Keywords: Synchronous Modeling, Ubiquitous Computing, middleware, internet of things

The aim of this research axis is to federate the inherent constraints of an activity recognition platform like SUP (see section 6.7) with a service-oriented middleware approach dealing with dynamic evolutions of system infrastructure in ubiquitous computing, and particularly in the Internet of Things (IoT). The Rainbow team (Nice-Sophia Antipolis University) proposes a component-based adaptive middleware (WComp [86], [85], [66]) to dynamically adapt and recompose assemblies of components.

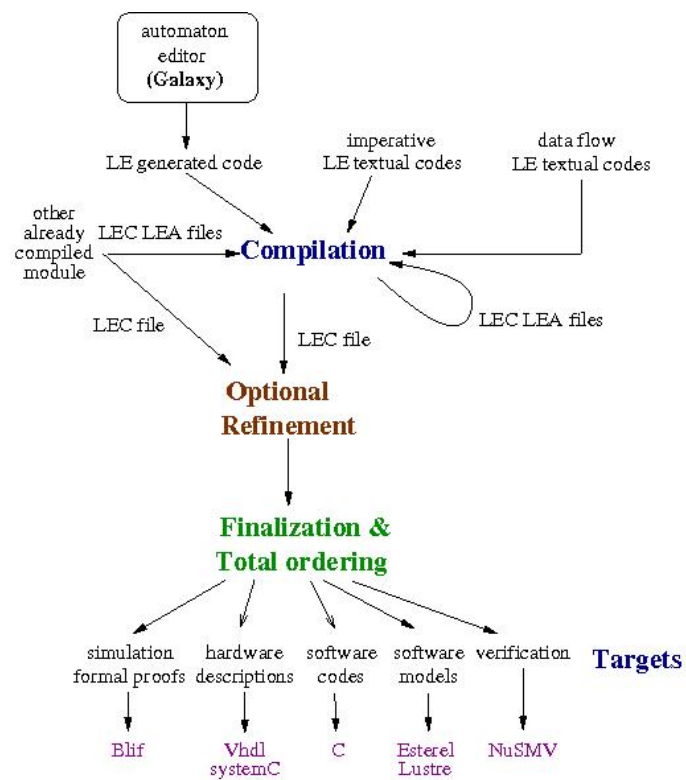


Figure 26. The Clem Toolkit

IoT is a way to combine computation and communication capabilities, sometimes in large scale information systems, with a huge number of complex devices connected to the physical world. Such infrastructures are often dedicated to the deployment of multiple applications, running concurrently. These applications are using shared devices from a common environment through different network middleware and numerous IoT protocols. Indeed, “Things”, also called the *Entities of Interest* [63], are the part of the real world in which devices are interacting and which must not be neglected. We aimed to model and validate concurrent accesses to shared devices without neglecting their associated *Entity of Interest*, their common physical context. One of the main challenge is then how to guarantee and validate some safety and integrity properties throughout the system’s evolution. In WComp middleware, we use synchronous models to facilitate the study and the validation of new composition mechanisms between applications at runtime. Then key problems to solve are: (1) how to specify and respect the "Thing" behavior? (2) how to ensure a safe combination of these multiple accesses when several services accesses a same entry of an *Entity of Interest* ? (3) how to manage multiple uses when applications simultaneously use a same service ?

This year, we addressed these problems by relying on formal method to model device behaviors as synchronous automata, taking into consideration their impact on the Entity Of Interest. Such an approach allows applying model-checking techniques to verify safety properties of applications. The main contribution is the definition of a sound way to compose models allowing context change adaptation. This composition relies on synchronous parallel composition paradigm. We proved that this operation preserves safety properties. However, it is not sufficient to obtain a global model of this composition because some devices may interact with the same Entity Of Interest. Moreover, several applications may use the same device services and then they can have concurrent accesses to their entries, so it can have an unexpected impact on our Entity Of Interest. Therefore, we added constraints to the device models composition and to applications level. We defined a generic way to express these constraints, independently of the knowledge about the devices and the applications, only their type is sufficient. We proposed the Description Constraint Language (DCL) to express these generic constraints and their compilation into LE Mealy machines. Thus we rely on CLEM model-checking facilities (see section 7.22) to validate the constraints. As a consequence, this approach ensures the adaptation to a context change and offers a means to formally perform validation.

These results have been published in [43]

7.24. Design of UHD Panoramic Video Camera

Participants: Carolina Da Silva Gomes Crispim, Rachid Guerchouche, Daniel Gaffé, François Brémond.

The goal of this work is to investigate the possibilities of designing a new camera-based system for retail. This work was carried in the context of a collaboration between STARS and Neosensys. The system is composed of several high definition cameras placed in a configuration such as it makes it possible to obtain a panoramic vision with 360° of field of view. The work was divided into 2 parts: theoretical part and practical part.

In the theoretical part, the different characteristics of the desired system were studied, such as: number of cameras, resolution of each camera, the different characteristics of the sensors (WDR, HDR, exposure) etc. Depending on these characteristics, data transmission through an IP network was addressed. In addition to the hardware characteristics, the possibility of embedding stitching capabilities was studied. After spending some time understanding the background behind such techniques, an existing implementation of the stitching was adopted. Simulations were then made in order to estimate the characteristics of an FPGA capable of handling 5 cameras with 12MP resolution each. An existing FPGA architecture extensively used in the industry was chosen and a mathematical model was provided in order to estimate the characteristics of such FPGA according to the different parameters of the camera-based system.

In the practical part, an implementation of the two first steps of the stitching algorithm (homograph estimation and warping) was performed on a FPGA using 2 cameras. The problems of code optimisation were addressed in order to achieve a functioning implementation with respect to the memory and computation capabilities of the FPGA.

7.25. Brick & Mortar Cookies

Participant: Anaïs Ducoffe.

The objective of the BMC project is to create a software that aims to present attendance and attractiveness of the customer in stores, based on automatic video analysis. This final system is designed to be used without changing current camera network of the customer store, dedicated to security purpose. Analysis should be given at different time and space resolutions. For instance, day attendance can be as interesting as year attendance. Moreover, shop owners want to be able to compare two given years or months, etc... As space resolution is concerned, the software should be able to give information about the global attractiveness of the store but should also analyze some specific zones.

IVA embedded on Bosch cameras

Intelligence Video Analysis (IVA) is embedded in some models of Bosch cameras. The algorithms are composed of human detection and tracking. They can be configured directly on the camera interface via *tasks*. Following Bosch *tasks* were selected and studied:

- **Loitering and idle object in a field tasks** enable to detect stop actions in a zone, when they happened and the stop positions.
- **Entering and leaving field tasks** enable to know when a person enters or leaves a zone.
- **Detect people in a field task** enables counting people in a zone.
- **Crossing lines tasks** for counting people entering or leaving shop. We are able to know when the line was crossed and in what sense.

It is not possible to get people trajectory when metadata from Bosch cameras are acquired in offline mode. Then we studied live connection to get metadata directly from the camera stream using a RTSP connection. Metadata information is saved in XML format.

The previously enumerated tasks use algorithms to detect people and get their trajectories. STARS team has developed similar algorithms and has adapted their parameters values to the specific needs of this software. Moreover these algorithms can be run on any type of video cameras (live and offline modes) whereas Bosch IVA can only be run in live mode on compatible Bosch cameras. Stars algorithms can also combine several cameras at the same time in order to track people across the camera network. We need those algorithms to sell a system that doesn't need a new camera network but can be used with existing ones. They will be integrated in the final product.

Tests in real conditions

A system for testing cameras and our software was installed in partner store (Super U). Cameras were installed and configured to process all our use cases and test our mechanism to extract the metadata. We used only Bosch camera with embedded IVA. We successfully acquire 2 hours of the desired metadata. The results of embedded algorithms are reliable on realistic data : we get good results in counting people and trajectories are accurate.

Metadata storage in database

Metadata have to be stored in a hierarchical way as request of the metadata by the application should be easy and quick. We choose to store metadata in a database. This database design was constraint by data storage speed and a quick access for live computation. Different parts of database (store information, devices description etc...) were designed to be as much independent as possible.

Web interface (GUI)

The graphic interface design is in progress. The interface will be a web based one to narrow compatibility problems: the application should be used as well with a computer as a tablet.

7.26. Monitoring Older People Experiments

Participants: Matias Marin, Etienne Corvée, François Brémond.

This year we have conducted many experiments, especially in Nice and partially in Thessaloniki, Lulea, Taiwan and Dublin, to validate our studies on monitoring older people suffering from various behavioral disorders in the framework of several projects.

DEM@CARE PROJECT

For the project Dem@care (see section 9.2.1.3), we use PCs with ASUS cameras, for monitoring and collecting a video dataset associated with metadata⁰. The software CAR is installed to automatically annotate the videos⁰. Data is recorded locally, and backups are made automatically and remotely: one on the server (LAB at the nursery home) and another backup at Inria. These data can be reached locally at the nursery home, thanks to the server located in the lab; also, they are all accessible from Inria network by ssh.

SafEE PROJECT

SafEE project (see section 9.1.1.2) experiments in the nursing home have started at the end of 2015, and are made up with different technologies (wifi, wired network, smart phones, Kinect, RFID, tablet...). In the nursing room, 2 PCs with KINECT2 are connected for monitoring the residents and are stored in a database. Another PC with Windows software is configured for SafEE serious game (cognitive games and music box, 1.7 version). Moreover, a Wifi access point will be used by medical staff at the nursery home to connect to a Graphical User Interface through a website designed by INDES team⁰, to consult patient data (their daily activities), from the activity history stored in the database or in real-time. Another device called AromaCare is installed in the rooms, which is a connected aroma diffuser by Wi-Fi. With the app Aroma Therapeutics (smartphone or tablet) we can manage several diffusers, by scheduling different programs each day and change the intensity of the diffusion.

In patients' home the same devices than Nursing Home have to be configured. Today, only recordings are done and stored at Inria.

OTHER PROJECTS at ICP (Institut Claude Pompidou)

ICP has now a remote access by using rdesktop, which is safer than team-viewer screen sharing session. The installation of new experimentations (e.g. praxis, relaxation, serious games) is now in progress: the configuration expected includes PCs with KINECT2 connected at ICP network and accessible from Inria. Some experimentations will use wireless sensors (e.g. accelerometer, pressure), controlled by the app wireless tag (on smartphone or tablet) to measure fine patient activities: motion, kettle utilization, etc.

⁰<https://team.inria.fr/stars/demcare-chu-dataset/>

⁰<https://team.inria.fr/stars/software/car-complex-activity-recognition-component-installation/>

⁰<http://webrobotics.inria.fr:8081/hop/events>

TITANE Project-Team

7. New Results

7.1. Analysis

7.1.1. Planar Shape Detection and Regularization in Tandem

Participants: Sven Oesau, Florent Lafarge, Pierre Alliez.

In collaboration with EADS ASTRIUM

We contributed a method for planar shape detection and regularization from raw point sets. The geometric modeling and processing of man-made environments from measurement data often relies upon robust detection of planar primitive shapes. In addition, the detection and reinforcement of regularities between planar parts is a means to increase resilience to missing or defect-laden data as well as to reduce the complexity of models and algorithms down the modeling pipeline. The main novelty behind our method is to perform detection and regularization in tandem. We first sample a sparse set of seeds uniformly on the input point set, then perform in parallel shape detection through region growing, interleaved with regularization through detection and reinforcement of regular relationships (coplanar, parallel and orthogonal). In addition to addressing the end goal of regularization, such reinforcement also improves data fitting and provides guidance for clustering small parts into larger planar parts (Figure 1). We evaluate our approach against a wide range of inputs and under four criteria: geometric fidelity, coverage, regularity and running times. Our approach compares well with available implementations such as the efficient RANSAC-based approach proposed by Schnabel and co-authors in 2007 [8]. This work has been published in the Computer Graphics Forum journal.

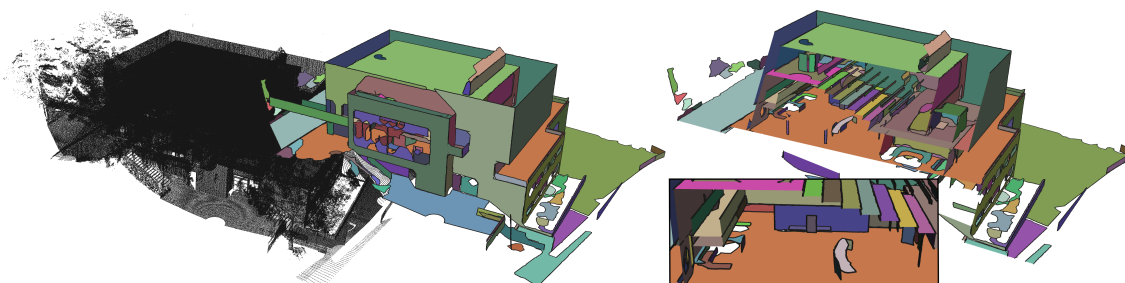


Figure 1. Shape detection and regularization. The input point set (5.2M points) has been acquired via a LIDAR scanner, from the inside and outside of a physical building. 200 shapes have been detected, aligned with 12 different directions in 179 different planes. The cross section depicts the auditorium in the upper floor and the entrance hall in the lower floor. The closeup highlights the steps of the auditorium which are made up of perfectly parallel and orthogonal planes.

7.1.2. Image partitioning into convex polygons

Participants: Liuyun Duan, Florent Lafarge.

In collaboration with Geoimage

The over-segmentation of images into atomic regions has become a standard and powerful tool in Vision. Traditional superpixel methods, that operate at the pixel level, cannot directly capture the geometric information disseminated into the images. We propose an alternative to these methods by operating at the level of geometric shapes. Our algorithm partitions images into convex polygons. It presents several interesting properties in terms of geometric guarantees, region compactness and scalability. The overall strategy consists in building a Voronoi diagram that conforms to preliminarily detected line-segments, before homogenizing the partition by spatial point process distributed over the image gradient. Our method is particularly adapted to images with strong geometric signatures, typically man-made objects and environments (Figure 2). We show the potential of our approach with experiments on large-scale images and comparisons with state-of-the-art superpixel methods [17]. This work has been published in the Computer Graphics Forum journal. Published in the proceedings of CVPR (IEEE conference on Computer Vision and Pattern Recognition).

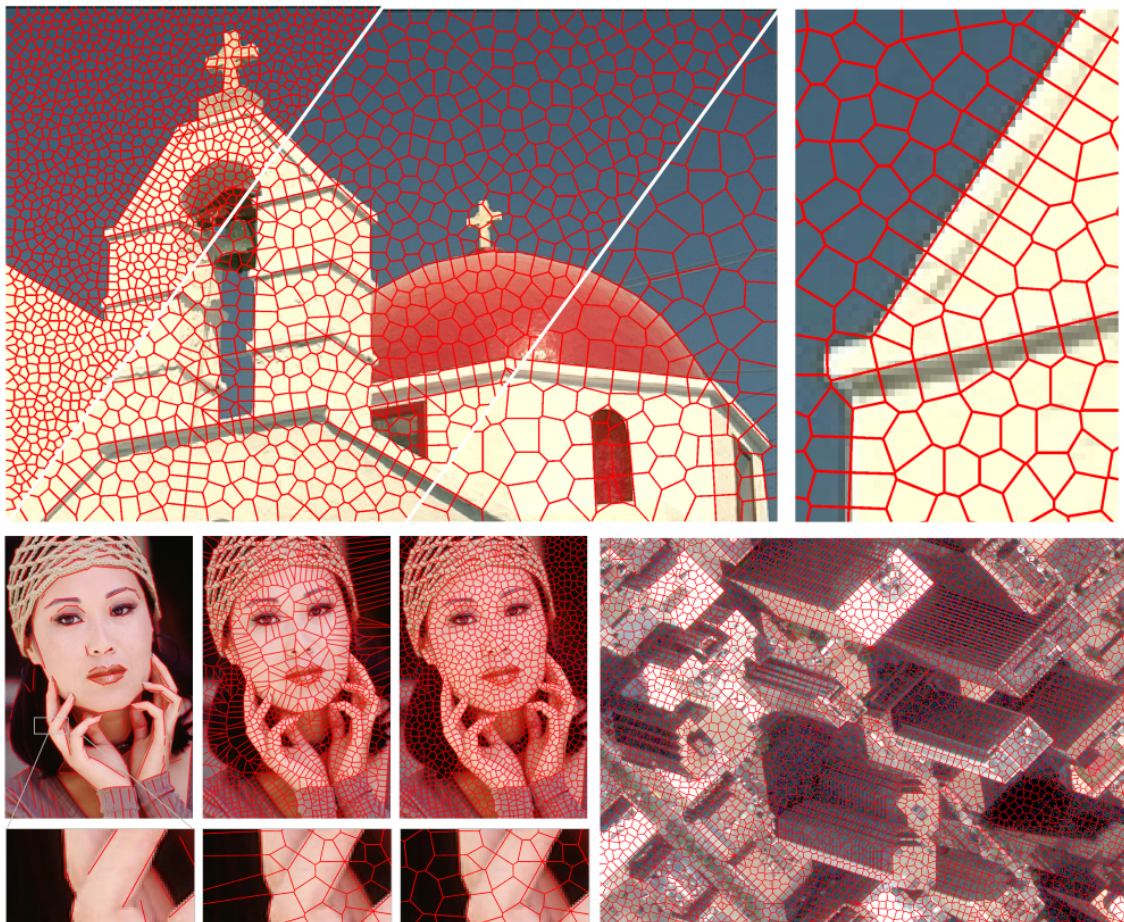


Figure 2. Image partitioning into convex polygons.

7.1.3. Object Classification via Planar Abstraction

Participants: Sven Oesau, Florent Lafarge, Pierre Alliez.

In collaboration with EADS ASTRIUM.

We contributed a supervised machine learning approach for classification of objects from sampled point data. The main idea consists in first abstracting the input object into planar parts at several scales, then discriminate between the different classes of objects solely through features derived from these planar shapes. Abstracting into planar shapes provides a means to both reduce the computational complexity and improve robustness to defects inherent to the acquisition process. Measuring statistical properties and relationships between planar shapes offers invariance to scale and orientation. A random forest is then used for solving the multiclass classification problem. We demonstrate the potential of our approach on a set of indoor objects from the Princeton shape benchmark and on objects acquired from indoor scenes and compare the performance of our method with other point-based shape descriptors [22] (see Figure 3).

7.1.4. Optimizing partition trees for multi-object segmentation with shape prior

Participants: Emmanuel Maggiori, Yuliya Tarabalka.

This work has been done in collaboration with Dr. Guillaume Charpiat (TAO team, Inria Saclay).

Partition trees, multi-class segmentation, shape priors, graph cut.

A partition tree is a hierarchical representation of an image. Once constructed, it can be repeatedly processed to extract information. Multi-object multi-class image segmentation with shape priors is one of the tasks that can be efficiently done upon an available tree. The traditional construction approach is a greedy clustering based on color similarities. However, not considering higher level cues during the construction phase leads to trees that might not accurately represent the underlying objects in the scene, inducing mistakes in the later segmentation. We proposed a method to optimize a tree based both on color distributions and shape priors [15]. It consists in pruning and regrafting tree branches in order to minimize the energy of the best segmentation that can be extracted from the tree. Theoretical guarantees help reduce the search space and make the optimization efficient. Our experiments (see Figure 4) show that we succeed in incorporating shape information to restructure a tree, which in turn enables to extract from it good quality multi-object segmentations with shape priors. Published in the proceedings of BMVC (British Machine Vision Conference).

7.2. Reconstruction

7.2.1. LOD Generation for Urban Scenes

Participants: Florent Lafarge, Pierre Alliez.

We contributed a novel approach that reconstructs 3D urban scenes in the form of levels of detail (LODs). Starting from raw data sets such as surface meshes generated by multi-view stereo systems, our algorithm proceeds in three main steps: classification, abstraction and reconstruction (Figure 5). From geometric attributes and a set of semantic rules combined with a Markov random field, we classify the scene into four meaningful classes. The abstraction step detects and regularizes planar structures on buildings, fits icons on trees, roofs and facades, and performs filtering and simplification for LOD generation. The abstracted data are then provided as input to the reconstruction step which generates watertight buildings through a min-cut formulation on a set of 3D arrangements. Our experiments on complex buildings and large scale urban scenes show that our approach generates meaningful LODs while being robust and scalable. By combining semantic segmentation and abstraction it also outperforms general mesh approximation approaches at preserving urban structures [10]. Published in the ACM Transactions on Graphics journal.

7.2.2. A Surface Reconstruction Method for In-Detail Underwater 3D Optical Mapping

Participant: Pierre Alliez.

In collaboration with Ricard Campos and Rafael Garcia from the Computer Vision and Robotics Group from University of Girona, and Mariette Yvinec from the GEOMETRICA Inria project-team.

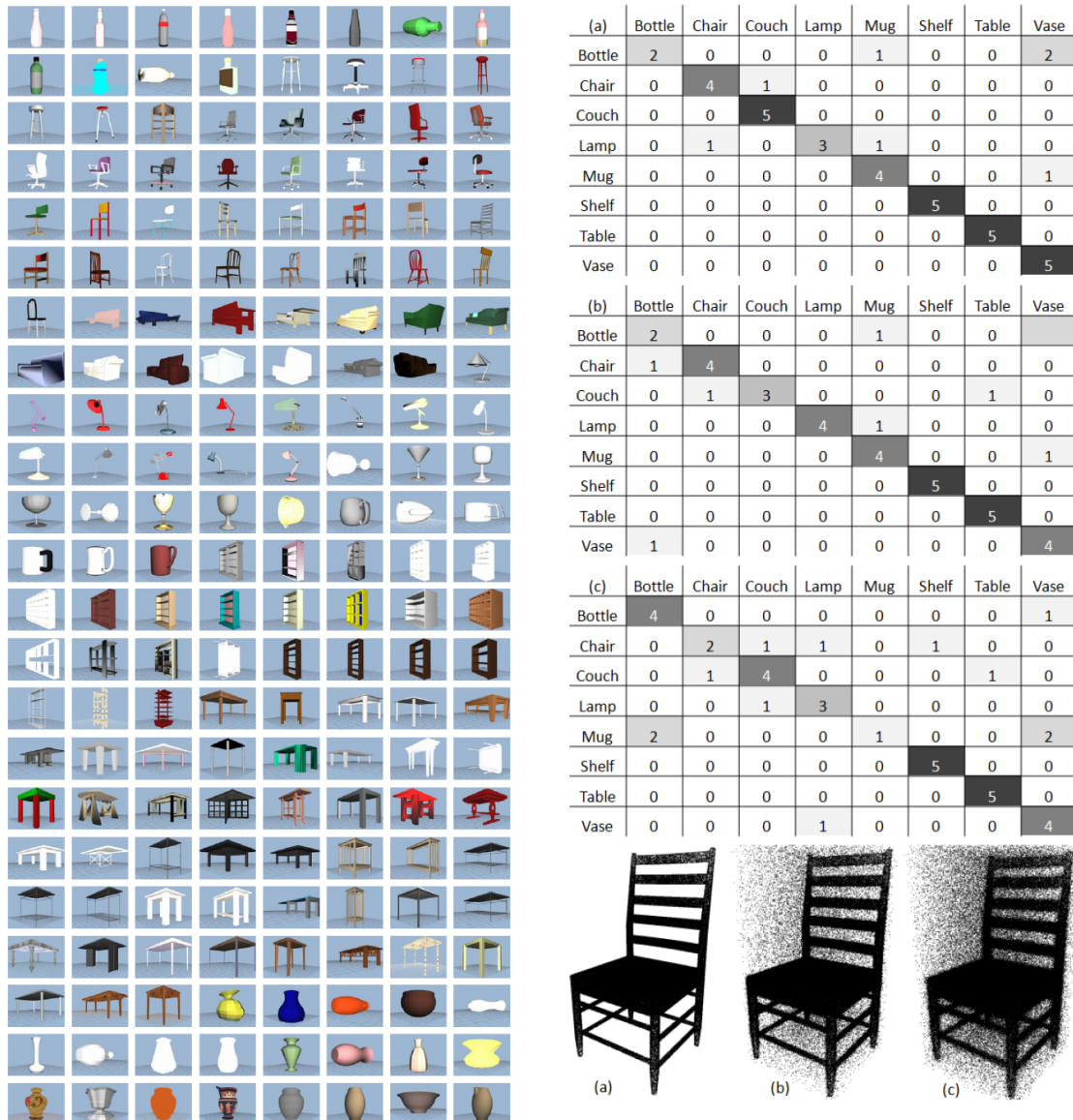


Figure 3. Classification. Left: We used four tabletop object classes from the Princeton Shape Benchmark: Bottle, Lamp, Mug and Vase. We also select four furniture object classes common to indoor scenes: Chair, Couch, Shelf and Table. Right: We evaluate our approach through computing a confusion matrix, for an increasing amount of noise and outliers. (a): Without noise and outliers. The precision of the class prediction is 82,5%. The classifier is not reliable for classifying the bottles, which get mislabeled as vases. (b): Added 10% outliers and 0.5% noise. Compared to the noise-free version the precision slightly dropped to 77.5%. (c): Added 20% outliers and 1% noise. The method maintains a precision of 70% for this level of noise.

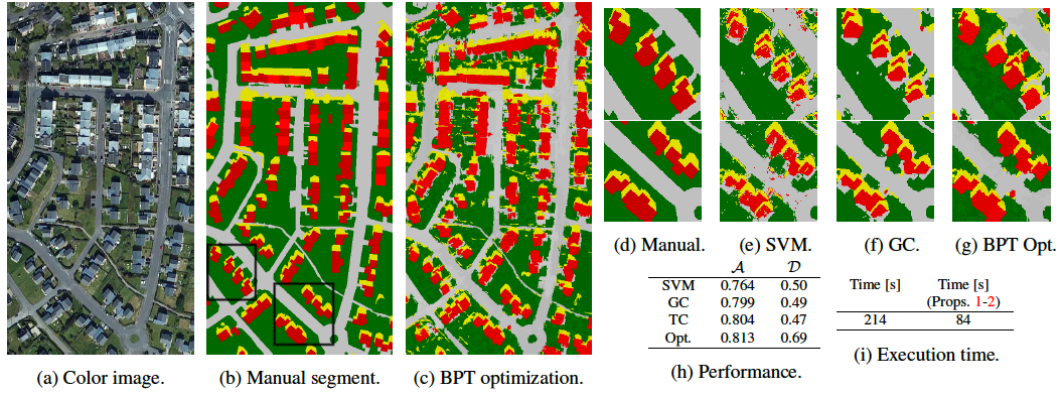


Figure 4. Classification results for the satellite image over Brest. \mathcal{A} denotes overall classification accuracy, and \mathcal{D} denotes average buildings overlap. The performance of the proposed binary partition tree (BPT) optimization method is compared with the following methods: 1) support vector machines (SVM) classification; 2) graph cut (GC) with α -expansion; 3) cut on the BPT, regularized by the number of regions without using shape priors (TC).

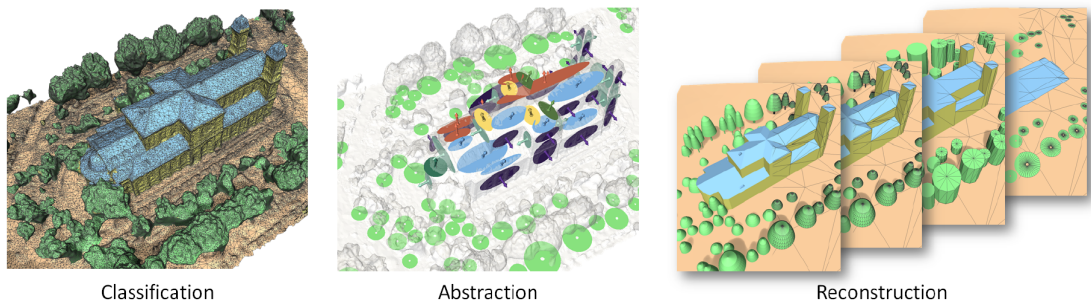


Figure 5. LOD Generation for Urban Scenes. Main steps of our algorithm.

Underwater range scanning techniques are starting to gain interest in underwater exploration, providing new tools to represent the seafloor. These scans (often) acquired by underwater robots usually result in an unstructured point cloud, but given the common downward-looking or forward-looking configuration of these sensors with respect to the scene, the problem of recovering a piecewise linear approximation representing the scene is normally solved by approximating these 3D points using a heightmap (2.5D). Nevertheless, this representation is not able to correctly represent complex structures, especially those presenting arbitrary concavities normally exhibited in underwater objects. We present a method devoted to full 3D surface reconstruction that does not assume any specific sensor configuration. The method presented is robust to common defects in raw scanned data such as outliers and noise often present in extreme environments such as underwater, both for sonar and optical surveys (Figure 6). Moreover, the proposed method does not need a manual preprocessing step. It is also generic as it does not need any information other than the points themselves to work. This property leads to its wide application to any kind of range scanning technologies and we demonstrate its versatility by using it on synthetic data, controlled laser-scans, and multibeam sonar surveys. Finally, and given the unbeatable level of detail that optical methods can provide, we analyze the application of this method on optical datasets related to biology, geology and archeology [4]. Published in the International Journal of Robotics Research.

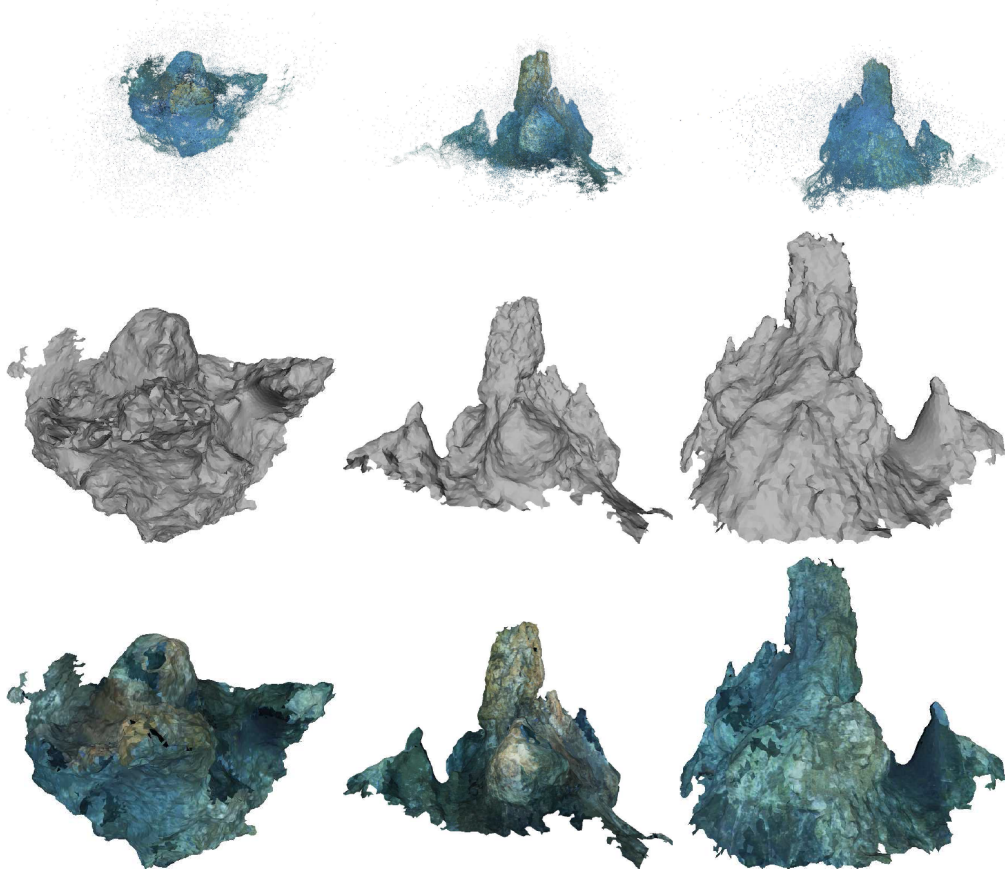


Figure 6. Underwater reconstruction.

7.2.3. Line Drawing Interpretation in a Multi-View Context

Participants: Jean-Dominique Favreau, Florent Lafarge.

In collaboration with Adrien Bousseau from the Inria project-team GraphDeco.

Many design tasks involve the creation of new objects in the context of an existing scene. Existing work in computer vision only provides partial support for such tasks. On the one hand, multi-view stereo algorithms allow the reconstruction of real-world scenes, while on the other hand algorithms for line-drawing interpretation do not take context into account. Our work combines the strength of these two domains to interpret line drawings of imaginary objects drawn over photographs of an existing scene. The main challenge we face is to identify the existing 3D structure that correlates with the line drawing while also allowing the creation of new structure that is not present in the real world. We propose a labeling algorithm to tackle this problem, where some of the labels capture dominant orientations of the real scene while a free label allows the discovery of new orientations in the imaginary scene (Figure 7). We illustrate our algorithm by interpreting line drawings for urban planning, home remodeling, furniture design and cultural heritage [18]. Published in the proceedings of CVPR (IEEE conference on Computer Vision and Pattern Recognition).

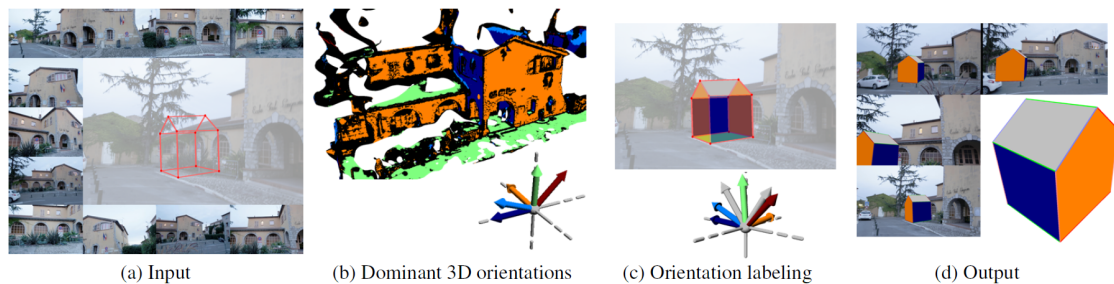


Figure 7. Line Drawing Interpretation. (a) Our algorithm takes as input multiple images of a scene along with a line-drawing traced over one of these images. (b) We first detect the dominant orientations of the existing scene from its multi-view stereo reconstruction. (c) Our labeling algorithm estimates the orientation of each facet of the drawing, favoring orientations already present in the scene. We visualize each dominant orientation with a random color; gray denotes new orientations. (d) We finally solve for the 3D model corresponding to the estimated orientations.

7.2.4. Marked point process model for curvilinear structures extraction

Participant: Yuliya Tarabalka [contact].

In collaboration with Seong-Gyun Jeong and Dr. Josiane Zerubia (AYIN team, Inria-SAM).

In this work, we proposed a new marked point process (MPP) model and the associated optimization technique to extract curvilinear structures [12]. Given an image, we compute the intensity variance and rotated gradient magnitude along the line segment. We constrain high level shape priors of the line segments to obtain smoothly connected line configuration. The optimization technique consists of two steps to reduce the significance of the parameter selection in our MPP model. We employ a Monte Carlo sampler with delayed rejection to collect line hypotheses over different parameter spaces. Then, we maximize the consensus among line detection results to reconstruct the most plausible curvilinear structures without parameter estimation process. Experimental results (see Figure 8) show that the algorithm effectively localizes curvilinear structures on a wide range of datasets.

7.2.5. Inference of curvilinear structure based on learning a ranking function and graph theory

Participant: Yuliya Tarabalka [contact].

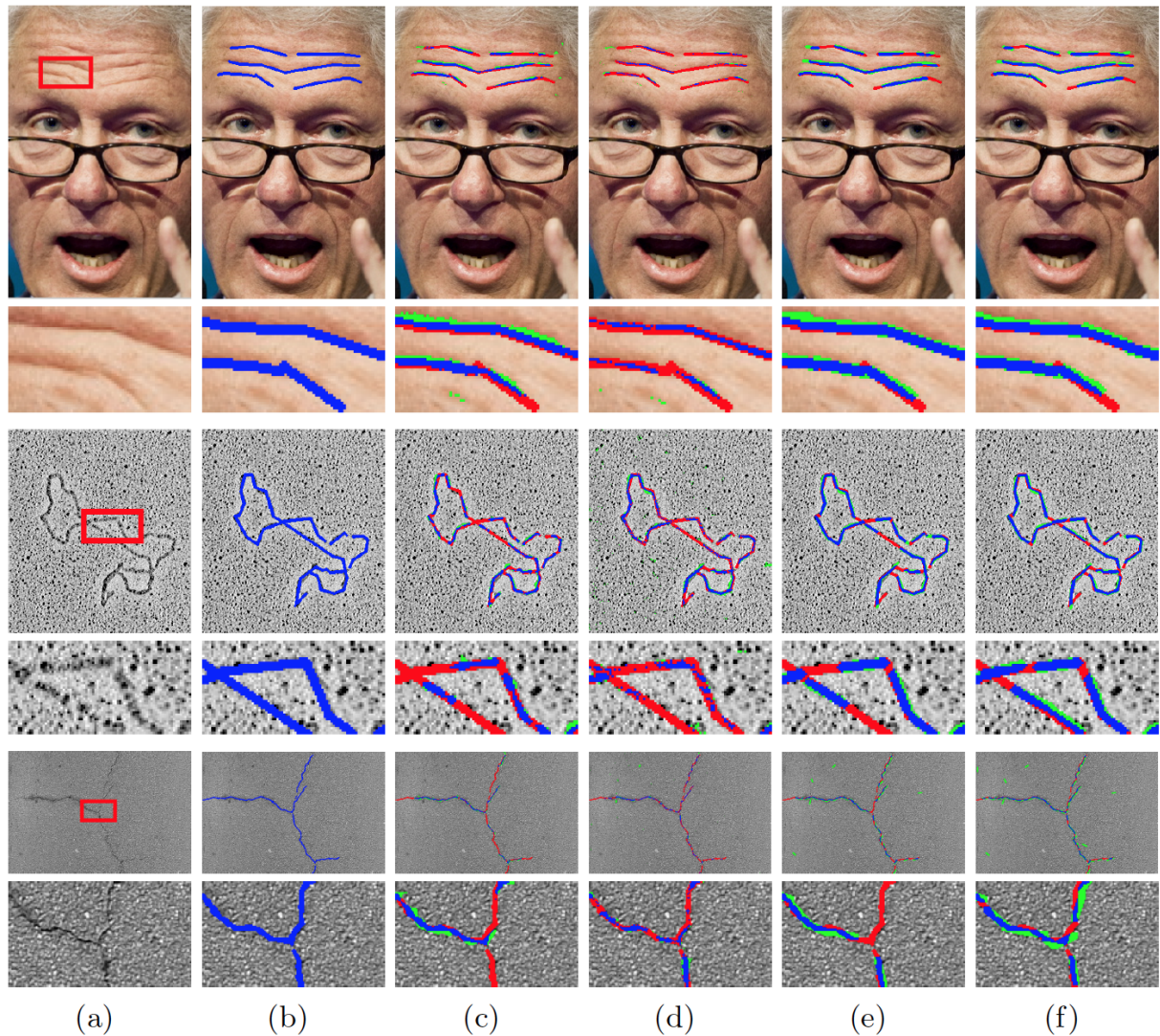


Figure 8. We visualize the localization of the curvilinear structures on input images (a). We compare with the results of a manually labeled image by a human expert (b), morphological filtering [Talbot 2007] (c), supervised feature learning [Becker 2013] (d), baseline MPP (e), and the proposed algorithm (f). Threshold values of (c) and (d) are chosen to achieve the closest recall scores to the proposed method. We use blue pixels to indicate areas which are completely corresponding to (b). Green and red pixels denote over-detected and under-detected areas, respectively, as compared with ground-truth. The name of the test images is from top to bottom: WRINKLE, DNA, and CRACK.

In collaboration with Seong-Gyun Jeong and Dr. Josiane Zerubia from the AYIN team and Dr. Nicolas Nisse from the COATI project-team.

Curvilinear structure extraction, inference of structured data, ranking learning, graphical model, shape simplification.

To detect curvilinear structures in natural images, we proposed a novel ranking learning system and an abstract curvilinear shape inference algorithm based on graph theory. We analyze the curvilinear structures as a set of small line segments. In this work, the rankings of the line segments are exploited to systematize the topological feature of the curvilinear structures. A Structured Support Vector Machine is employed to learn the ranking function that predicts the correspondence of the given line segments and the latent curvilinear structures. We first extract curvilinear features using morphological profiles and steerable filter responses. Also, we propose an orientation-aware feature descriptor and a feature grouping operator to improve the structural integrity during the learning process. To infer the curvilinear structure, we build a graph based on the output rankings of the line segments. We progressively reconstruct the curvilinear structure by looking for paths between remote vertices in the graph. Experimental results (see Figure 9 for an example of the experimental results' comparison on the CRACK dataset) show that the proposed algorithm faithfully detects the curvilinear structures within various datasets.

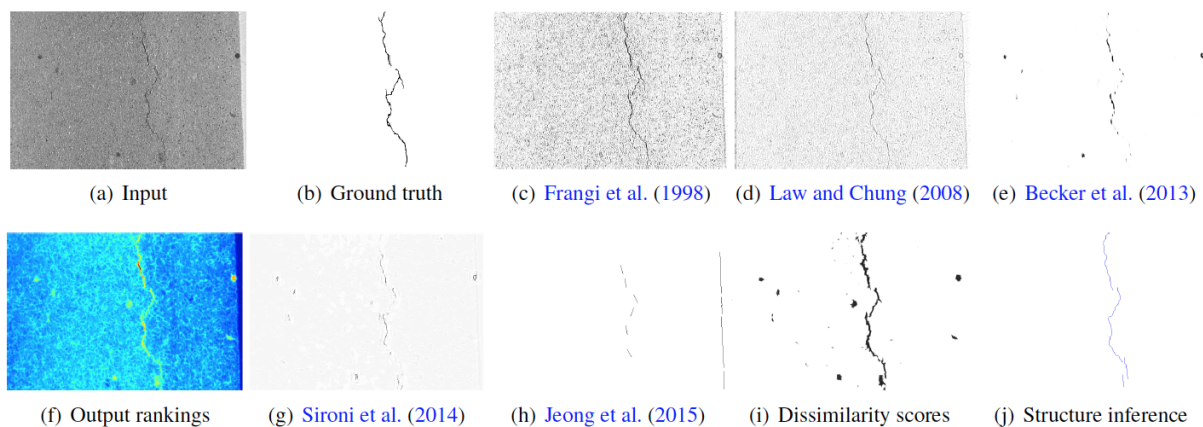


Figure 9. Inference of curvilinear structure on the CRACK dataset. Our approach is depicted by (j).

7.3. Approximation

7.3.1. Isotropic approximation within a tolerance volume

Participants: Manish Mandad, Pierre Alliez.

In collaboration with David Cohen-Steiner from the GEOMETRICA project-team.

We introduce an algorithm that generates a surface triangle mesh given an input tolerance volume. The mesh is guaranteed to be within the tolerance, intersection free and topologically correct. A pliant meshing algorithm is used to capture the topology and discover the anisotropy in the input tolerance volume in order to generate a concise output. We first refine a 3D Delaunay triangulation over the tolerance volume while maintaining a piecewise-linear function on this triangulation, until an isosurface of this function matches the topology sought after. We then embed the isosurface into the 3D triangulation via mutual tessellation, and simplify it while preserving the topology. Our approach extends to surfaces with boundaries and to non-manifold surfaces. We demonstrate the versatility of our approach on a variety of data sets and tolerance volumes [7]. Figure 10 illustrates the robustness of our approach on defect-laden inputs.

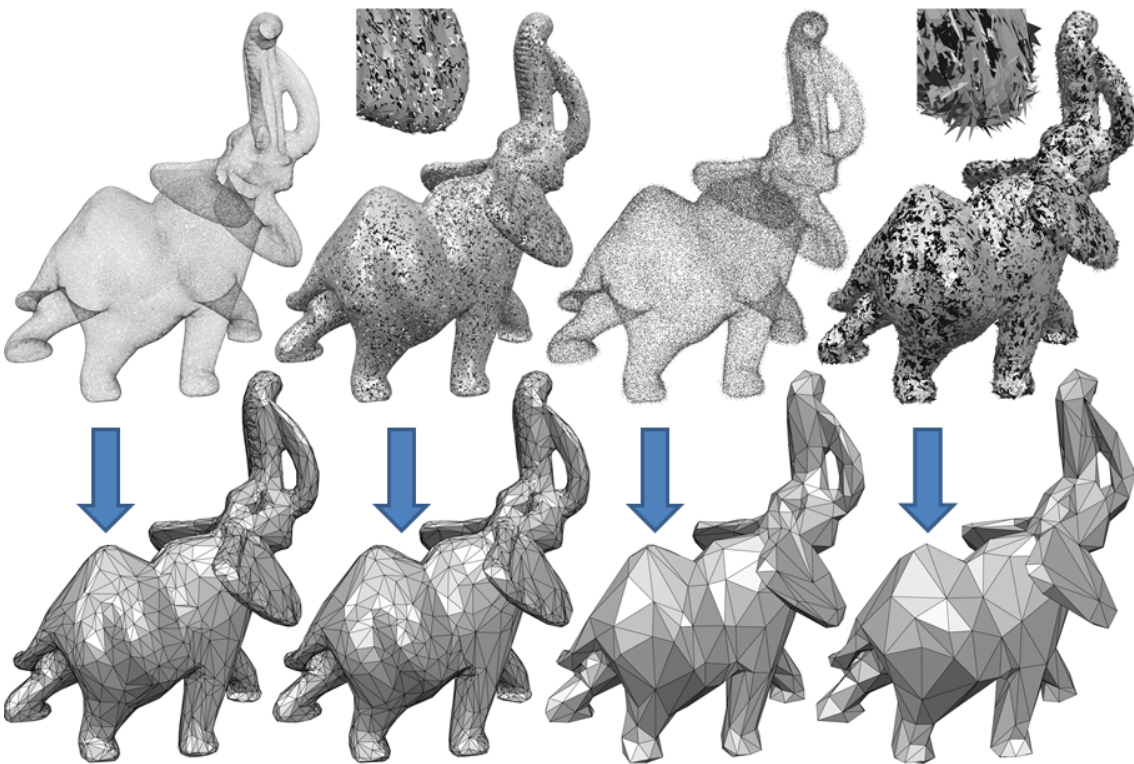


Figure 10. Isotropic approximation within a tolerance volume.

7.3.2. Structure-Aware Mesh Decimation

Participants: David Salinas, Florent Lafarge, Pierre Alliez.

We contributed to a novel approach for the decimation of triangle surface meshes. Our algorithm takes as input a triangle surface mesh and a set of planar proxies detected in a pre-processing analysis step, and structured via an adjacency graph. It then performs greedy mesh decimation through a series of edge collapse operators, designed to approximate the local mesh geometry as well as the geometry and structure of proxies (Figure 11). Such structure-preserving approach is well suited to planar abstraction, i.e., extreme decimation approximating well the planar parts while filtering out the others. Our experiments on a variety of inputs illustrate the potential of our approach in terms of improved accuracy and preservation of structure [9].

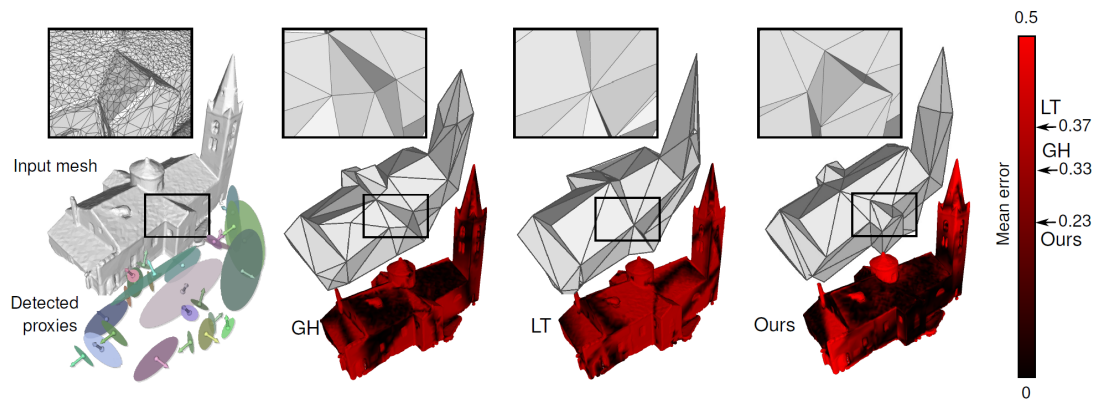


Figure 11. Structure-aware mesh decimation. Our algorithm simplifies dense triangle surface meshes via a structured set of planar proxies (left) that guides the decimation process while preserving the structure. At coarse complexity (here 50 vertices), common mesh decimation approaches (middle, quadric error metrics from Garland-Heckbert, and volume-preserving from Lindstrom-Turk) fail to reach low approximation error (see colored meshes) while preserving structure (see closeups).

7.3.3. CGALmesh: a Generic Framework for Delaunay Mesh Generation

Participants: Clément Jamin, Pierre Alliez.

In collaboration with Mariette Yvinec and Jean-Daniel Boissonnat from the GEOMETRICA project-team.

CGALmesh is the mesh generation software package of the Computational Geometry Algorithm Library (CGAL). It generates isotropic simplicial meshes – surface triangular meshes or volume tetrahedral meshes – from input surfaces, 3D domains as well as 3D multi-domains, with or without sharp features (see Figure 12). The underlying meshing algorithm relies on restricted Delaunay triangulations to approximate domains and surfaces, and on Delaunay refinement to ensure both approximation accuracy and mesh quality. CGALmesh provides guarantees on approximation quality as well as on the size and shape of the mesh elements. It provides four optional mesh optimization algorithms to further improve the mesh quality. A distinctive property of CGALmesh is its high flexibility with respect to the input domain representation. Such a flexibility is achieved through a careful software design, gathering into a single abstract concept, denoted by the oracle, all required interface features between the meshing engine and the input domain. We already provide oracles for domains defined by polyhedral and implicit surfaces [6].

7.4. Watermarking

7.4.1. Anti-Cropping Blind Resynchronization for 3D Watermarking

Participants: Xavier Rolland-Nevière, Pierre Alliez, Gwenaël Doërr.

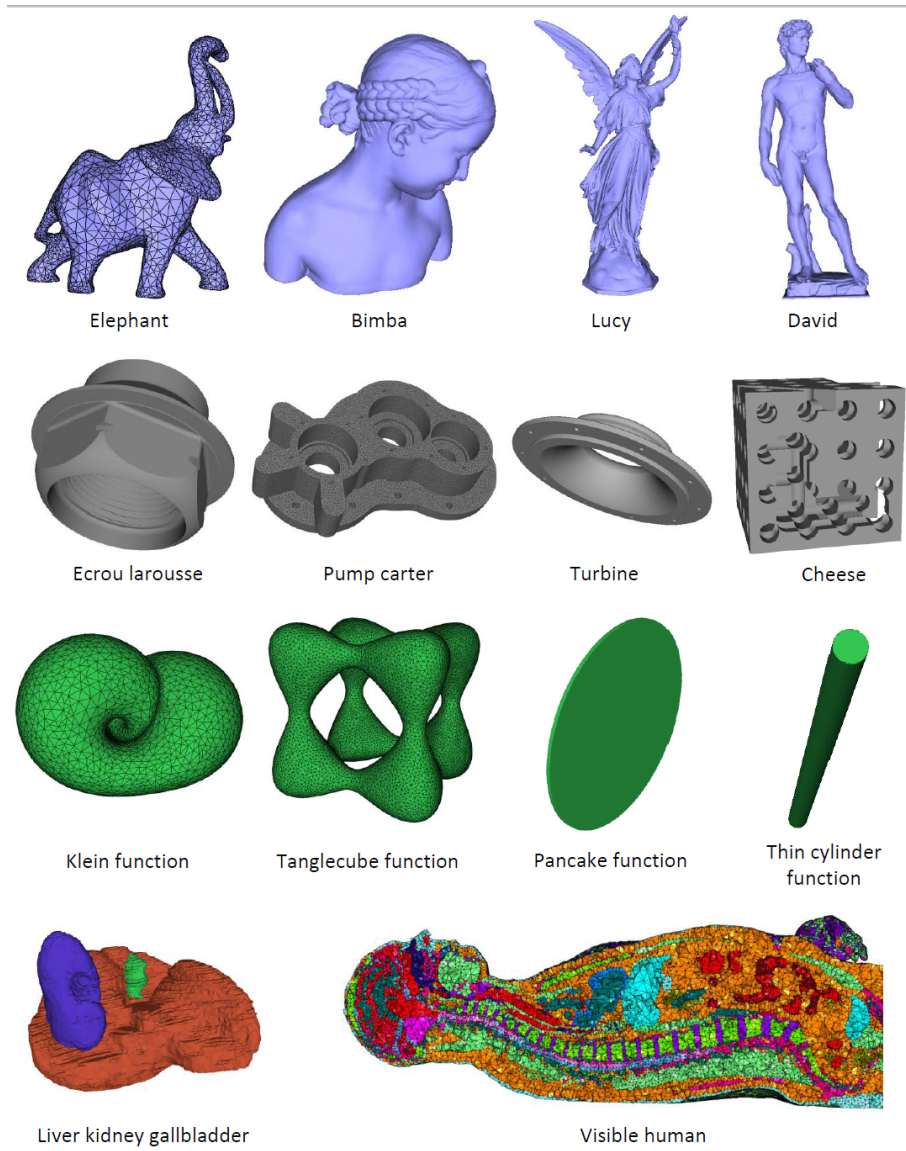


Figure 12. Input domains: domains bounded by smooth surfaces (top row, in blue), CAD models with sharp features (second row, in grey), implicit functions (third row, in green), 3D images (bottom row, multicolored).

Radial-based 3D watermarking alters the distances between the center of mass of the 3D mesh and its vertices. These watermarking systems are inherently sensitive to cropping. To address this limitation, this paper introduces a complementary blind resynchronization module to transmit critical synchronization information to the watermark decoder. Spherical patterns formed by several secret landmark vertices are embedded alongside the payload and blindly retrieved by the decoder, thereby conveying the synchronization information needed. Experimental results showcase significant improvement against cropping, while preserving performances against volumetric attacks thanks to a control parameter that automatically switches between alternate resynchronization modes [16].

TYREX Project-Team

6. New Results

6.1. Expressive Logical Combinators

A popular technique for the analysis of web query languages relies on the translation of queries into logical formulas. These formulas are then solved for satisfiability using an off-the-shelf satisfiability solver. A critical aspect in this approach is the size of the obtained logical formula, since it constitutes a factor that affects the combined complexity of the global approach. In this work [21], we present logical combinators whose benefit is to provide an exponential gain in succinctness in terms of the size of the logical representation. This opens the way for solving a wide range of problems such as satisfiability and containment for expressive query languages in exponential-time, even though their direct formulation into the underlying logic results in an exponential blowup of the formula size, yielding an incorrectly presumed two-exponential time complexity. We illustrate this from a practical point of view on a few examples such as numerical occurrence constraints and tree frontier properties which are concrete problems found with semi-structured data [21].

6.2. Behavioural Types

Behavioural type systems ensure more than the usual safety guarantees of static analysis. They are based on the idea of “types-as-processes”, providing dedicated type algebras for particular properties, ranging from protocol compatibility to race-freedom, lock-freedom, or even responsiveness.

Two successful, although rather different, approaches, are session types and process types. The former allows to specify and verify (distributed) communication protocols using specific type (proof) systems; the latter allows to infer from a system specification a process abstraction on which it is simpler to verify properties, using a generic type (proof) system. What is the relationship between these approaches? Can the generic one subsume the specific one? At what price? And can the former be used as a compiler for the latter?

In [15], we showed how communication protocols can be integrated into an object-oriented type system supporting *non-uniform objects*, i. e. objects where the sequences of method calls are restricted, such as a `File` where `read()` cannot be called after `close()`. In such a system, communication protocols can be enforced by giving appropriate non-uniform types to the socket objects. We defined a sound and complete type checking algorithm for a small distributed class-based object-oriented language with structural subtyping. Static typing guarantees that both sequences of messages on channels, and sequences of method calls on objects, conform to type-theoretic specifications, thus ensuring type-safety.

6.3. SPARQL Queries

Static analysis is a core task in query optimization and knowledge base verification. In [14], [24], we study static analysis techniques for SPARQL, the standard language for querying Semantic Web data. We are interested in developing techniques through reductions to the validity problem in logic

In [22], we investigate techniques for detecting SPARQL query update independence. A query is independent of an update when the execution of the update does not affect the result of the query. Determining independence is especially useful in the context of huge RDF repositories, where it permits to avoid expensive yet useless re-evaluation of queries. While this problem has been intensively studied for fragments of relational calculus, very few works exist for the standard query language for the semantic web. We report on our investigations on how a notion of independence can be defined in the SPARQL context.

6.4. Semantic Subtyping

In a programming language, subtyping represents a notion of safe substitutability (it is always safe to replace a value of some type with a value of a subtype). There are several ways such a relation can be formally defined. Semantic subtyping consists of giving a set-theoretic denotation to types and using set inclusion to define subtyping. Works by Benzaken, Castagna, Frisch and Xu have described how to define such relations for complex type algebras comprising recursive, product, function, intersection, union, and complement types together with type variables. In [17], we showed how to formalise such a relation in logic and decide it in EXPTIME, answering an open question, and discussed experiments made with the full implementation of the system in our solver (5.3).

6.5. Spatio-temporal validation of multimedia documents

A multimedia document authoring system should provide analysis and validation tools that help authors find and correct mistakes before document deployment. Although very useful, multimedia validation tools are not often provided. Spatial validation of multimedia documents may be performed over the initial position of media items before presentation starts. However, such an approach does not lead to good results when media item placement changes over time. Some document authoring languages allow the definition of spatio-temporal relationships among media items and they can be moved or resized during runtime. Current validation approaches do not verify dynamic spatio-temporal relationships. In [19], we present a novel approach for spatio-temporal validation of multimedia documents. We model the document state, extending the Simple Hypermedia Model (SHM), comprising media item positioning during the whole document presentation. Mappings between document states represent time lapse or user interaction. We also define a set of atomic formulas upon which the author's expectations related to the spatio-temporal layout can be described and analyzed.

6.6. XQuery and Static Typing

XQuery is a functional language dedicated to XML data querying and manipulation. As opposed to other W3C-standardized languages for XML (e.g. XSLT), it has been intended to feature strong static typing. Currently, however, some expressions of the language cannot be statically typed with any precision.

In [20], we argue that this is due to a discrepancy between the semantics of the language and its type algebra. We discuss how to handle this discrepancy by improving the type system. We describe a logic-based language of extended types able to represent inner tree nodes and show how it can dramatically increase the precision of typing for navigation expressions. We describe how inclusion between these extended types and the classical regular tree types can be decided, allowing a hybrid system combining both type languages. The result is a net increase in precision of typing.

In a previous work, we aimed at bridging the gap between path-based XML processing languages like XQuery and pattern-based such languages like CDuce. We extend the language CDuce into a succinct core λ -calculus that captures XQuery 3.0. The extensions we consider essentially allow CDuce to implement XPath-like navigational expressions by pattern matching and precisely type them. The elaboration of XQuery 3.0 into the extended CDuce provides a formal semantics and a sound static type system for XQuery 3.0 programs.

6.7. Efficiently Deciding μ -calculus with Converse over Finite Trees

In [16], we present a sound and complete satisfiability-testing algorithm and its effective implementation for an alternation-free modal μ -calculus with converse, where formulas are cycle-free and are interpreted over finite ordered trees. The time complexity of the satisfiability-testing algorithm is $2^O(n)$ in terms of formula size n . The algorithm is implemented using symbolic techniques (BDD). We present crucial implementation techniques and heuristics that we used to make the algorithm as fast as possible in practice. Our implementation is detailed in 5.3.

6.8. Reasoning with Style

The Cascading Style Sheets (CSS) language constitutes a key component of web applications. It offers a series of sophisticated features to stylize web pages. Its apparent simplicity and power are however counterbalanced by the difficulty of debugging and maintaining style sheets, tasks for which developers still lack appropriate tools. In particular, significant portions of CSS code become either useless or redundant, and tend to accumulate over time. The situation becomes even worse as more complex features are added to the CSS language (e.g. CSS3 powerful selectors). A direct consequence is a waste of CPU that is required to display web pages, as well as the significant amount of useless traffic at web scale. Style sheets are designed to operate on a set of documents (possibly generated). However, existing techniques consist in syntax validators, optimizers and runtime debuggers that operate in one particular document instance. As such, they do not provide guarantees concerning all web pages in CSS refactoring, such as preservation of the formatting. This is partly because they are essentially syntactic and do not take advantage of CSS semantics to detect redundancies. In [18], we propose a set of automated refactoring techniques aimed at removing redundant and inaccessible declarations and rules, without affecting the layout of any document to which the style sheet is applied. We implemented a prototype that has been extensively tested with popular web sites (such as Google Sites, CNN, Apple, etc.). We show that significant size reduction can be obtained while preserving the code readability and improving maintainability.

6.9. A Comparative Analysis of Attitude Estimation

We investigate the precision of attitude estimation techniques in the context of pedestrian dead-reckoning with commodity smartphones. We propose a comparative analysis of state-of-the-art algorithms for attitude estimation in this setting. We provide an experimental setup with a precise ground truth obtained with a motion capture system. We precisely quantify the error in attitude estimation obtained with each technique. We discuss the obtained results and analyse advantages and limitations of current technology for further PDR research.

WILLOW Project-Team

7. New Results

7.1. 3D object and scene modeling, analysis, and retrieval

7.1.1. *The joint image handbook*

Participants: Matthew Trager, Martial Hebert, Jean Ponce.

Given multiple perspective photographs, point correspondences form the “joint image”, effectively a replica of three-dimensional space distributed across its two-dimensional projections. This set can be characterized by multilinear equations over image coordinates, such as epipolar and trifocal constraints. In this work, we revisit the geometric and algebraic properties of the joint image, and address fundamental questions such as how many and which multilinearities are necessary and/or sufficient to determine camera geometry and/or image correspondences. Our new theoretical results answer these questions in a very general setting, and our work, published ICCV 2015 [17], is intended to serve as a “handbook” reference about multilinearities for practitioners.

7.1.2. *Trinocular Geometry Revisited*

Participants: Jean Ponce, Martial Hebert, Matthew Trager.

When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. In this project, we use elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes, which can be used to construct a practical and efficient method for trinocular geometry parameter estimation. This work has been published at CVPR 2014, and a revised version that includes numerical experiments using synthetic and real data has been submitted to IJCV [25].

7.1.3. *24/7 place recognition by view synthesis*

Participants: Akihiko Torii, Relja Arandjelović, Josef Sivic, Masatoshi Okutomi, Tomas Pajdla.

We address the problem of large-scale visual place recognition for situations where the scene undergoes a major change in appearance, for example, due to illumination (day/night), change of seasons, aging, or structural modifications over time such as buildings built or destroyed. Such situations represent a major challenge for current large-scale place recognition methods. This work has the following three principal contributions. First, we demonstrate that matching across large changes in the scene appearance becomes much easier when both the query image and the database image depict the scene from approximately the same viewpoint. Second, based on this observation, we develop a new place recognition approach that combines (i) an efficient synthesis of novel views with (ii) a compact indexable image representation. Third, we introduce a new challenging dataset of 1,125 camera-phone query images of Tokyo that contain major changes in illumination (day, sunset, night) as well as structural changes in the scene. We demonstrate that the proposed approach significantly outperforms other large-scale place recognition techniques on this challenging data. This work has been published at CVPR 2015 [16]. Figure 1 shows examples of the newly collected Tokyo 24/7 dataset.

7.1.4. *NetVLAD: CNN architecture for weakly supervised place recognition*

Participants: Relja Arandjelović, Petr Gronat, Akihiko Torii, Tomas Pajdla, Josef Sivic.



Figure 1. Example query images from the newly collected 24/7 Tokyo dataset. Each place in the query set is captured at different times of day: (a) daytime, (b) sunset, and (c) night. For comparison, the database street-view image at a close-by position is shown in (d). Note the major changes in appearance (illumination changes in the scene) between the database image (d) and the query images (a,b,c)

In [21], we tackle the problem of large scale visual place recognition, where the task is to quickly and accurately recognize the location of a given query photograph. We present the following three principal contributions. First, we develop a convolutional neural network (CNN) architecture that is trainable in an end-to-end manner directly for the place recognition task. The main component of this architecture, NetVLAD, is a new generalized VLAD layer, inspired by the "Vector of Locally Aggregated Descriptors" image representation commonly used in image retrieval. The layer is readily pluggable into any CNN architecture and amenable to training via backpropagation. Second, we develop a training procedure, based on a new weakly supervised ranking loss, to learn parameters of the architecture in an end-to-end manner from images depicting the same places over time downloaded from Google Street View Time Machine. Finally, we show that the proposed architecture obtains a large improvement in performance over non-learned image representations as well as significantly outperforms off-the-shelf CNN descriptors on two challenging place recognition benchmarks. This work is under review. Figure 2 shows some qualitative results.

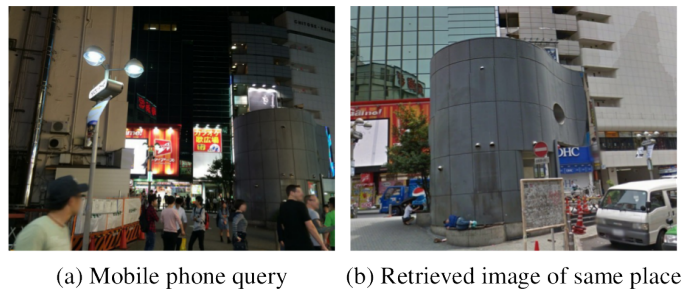


Figure 2. Our trained NetVLAD descriptor correctly recognizes the location (b) of the query photograph (a) despite the large amount of clutter (people, cars), changes in viewpoint and completely different illumination (night vs daytime).

7.2. Category-level object and scene recognition

7.2.1. Is object localization for free? – Weakly-supervised learning with convolutional neural networks

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

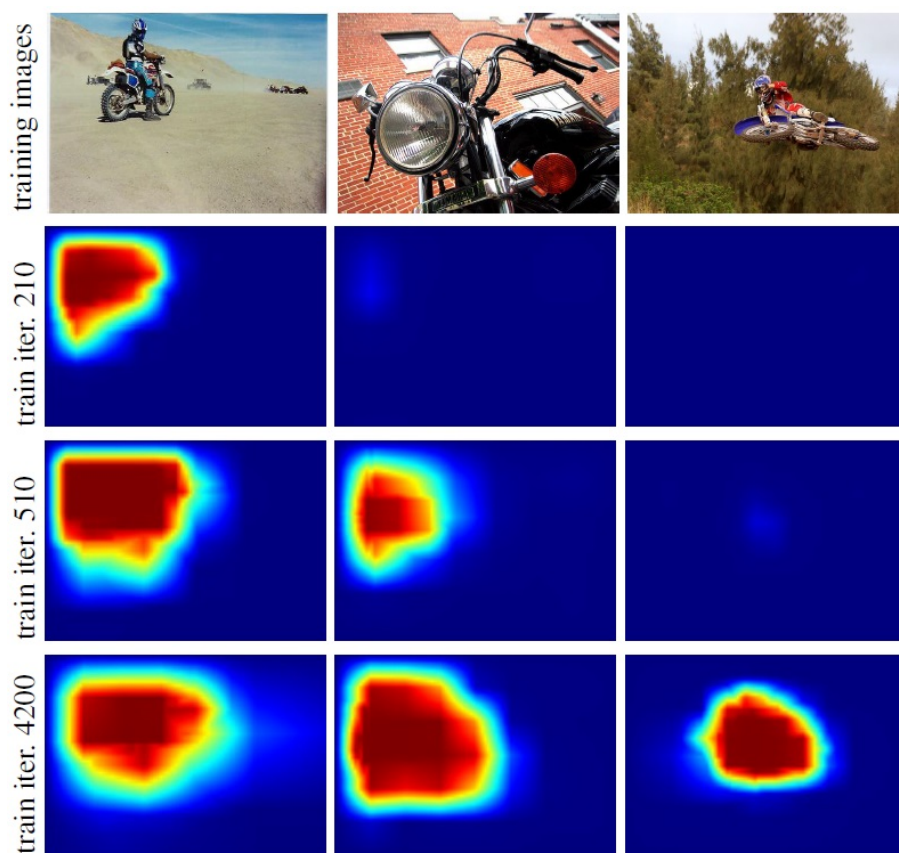


Figure 3. Evolution of localization score maps for the motorbike class over iterations of our weakly-supervised CNN training. Note that locations of objects with more usual appearance are discovered earlier during training.

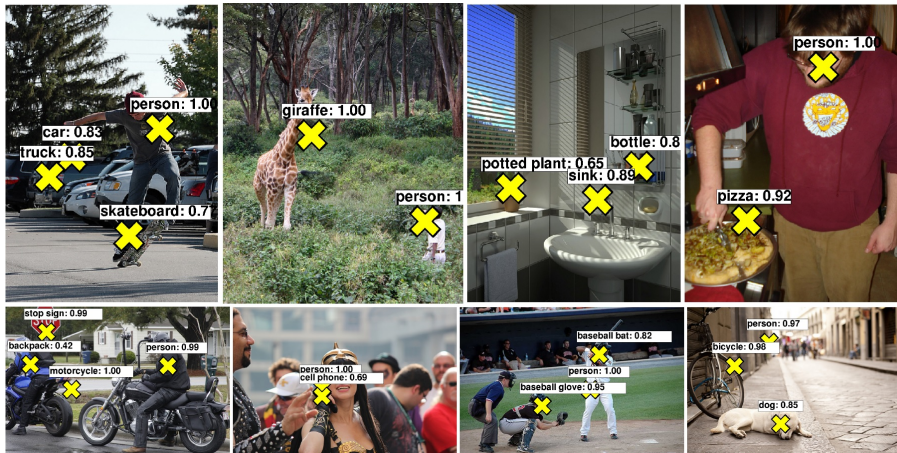


Figure 4. Example location predictions for images from the Microsoft COCO validation set obtained by our weakly-supervised method. Note that our method does not use object locations at training time, yet can predict locations of objects in test images (yellow crosses). The method outputs the most confident location for most confident object classes.

Successful methods for visual object recognition typically rely on training datasets containing lots of richly annotated images. Detailed image annotation, e.g. by object bounding boxes, however, is both expensive and often subjective. We describe a weakly supervised convolutional neural network (CNN) for object classification that relies only on image-level labels, yet can learn from cluttered scenes containing multiple objects (see Figure 3). We quantify its object classification and object location prediction performance on the Pascal VOC 2012 (20 object classes) and the much larger Microsoft COCO (80 object classes) datasets. We find that the network (i) outputs accurate image-level labels, (ii) predicts approximate locations (but not extents) of objects, and (iii) performs comparably to its fully-supervised counterparts using object bounding box annotation for training. This work has been published at CVPR 2015 [14]. Illustration of localization results by our method in Microsoft COCO dataset is shown in Figure 4.

7.2.2. Unsupervised Object Discovery and Localization in the Wild: Part-based Matching with Bottom-up Region Proposals

Participants: Minsu Cho, Suha Kwak, Cordelia Schmid, Jean Ponce.

In [8], we address *unsupervised* discovery and localization of dominant objects from a noisy image collection of multiple object classes. The setting of this problem is fully unsupervised (Fig. 5), without even image-level annotations or any assumption of a single dominant class. This is significantly more general than typical colocalization, cosegmentation, or weakly-supervised localization tasks. We tackle the unsupervised discovery and localization problem using a part-based region matching approach: We use off-the-shelf region proposals to form a set of candidate bounding boxes for *objects* and *object parts*. These regions are efficiently matched across images using a probabilistic Hough transform that evaluates the confidence for each candidate correspondence considering both appearance similarity and spatial consistency. Dominant objects are discovered and localized by comparing the scores of candidate regions and selecting those that stand out over other regions containing them. Extensive evaluations on standard benchmarks (e.g., Object Discovery and PASCAL VOC 2007 datasets) demonstrate that the proposed approach significantly outperforms the current state of the art in colocalization, and achieves robust object discovery even in a fully unsupervised setting. This work has been published in CVPR 2015 [8] as oral presentation.

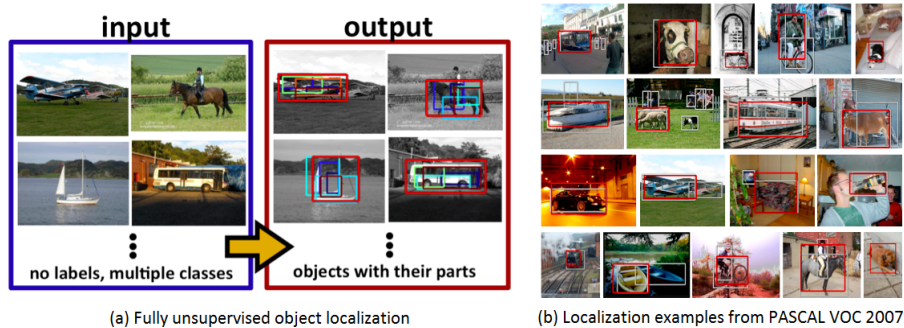


Figure 5. Unsupervised object discovery in the wild. We tackle object localization in an unsupervised scenario without any type of annotations, where a given image collection may contain multiple dominant object classes and even outlier images. The proposed method discovers object instances (red bounding boxes) with their distinctive parts (smaller boxes).

7.2.3. Unsupervised Object Discovery and Tracking in Video Collections

Participants: Suha Kwak, Minsu Cho, Ivan Laptev, Jean Ponce, Cordelia Schmid.

In [11], we address the problem of automatically localizing dominant objects as spatio-temporal tubes in a noisy collection of videos with minimal or even no supervision. We formulate the problem as a combination of two complementary processes: discovery and tracking (Figure 6). The first one establishes correspondences between prominent regions across videos, and the second one associates similar object regions within the same video. It is empirically demonstrated that our method can handle video collections featuring multiple object classes, and substantially outperforms the state of the art in colocalization, even though it tackles a broader problem with much less supervision. This work has been published in ICCV 2015.

7.2.4. Linking Past to Present: Discovering Style in Two Centuries of Architecture

Participants: Stefan Lee, Nicolas Maisonneuve, David Crandall, Alexei A. Efros, Josef Sivic.

With vast quantities of imagery now available online, researchers have begun to explore whether visual patterns can be discovered automatically. Here we consider the particular domain of architecture, using huge collections of street-level imagery to find visual patterns that correspond to semantic-level architectural elements distinctive to particular time periods. We use this analysis both to date buildings, as well as to discover how functionally similar architectural elements (e.g. windows, doors, balconies, etc.) have changed over time due to evolving styles. We validate the methods by combining a large dataset of nearly 150,000 Google Street View images from Paris with a cadastre map to infer approximate construction date for each facade. Not only could our analysis be used for dating or geo-localizing buildings based on architectural features, but it also could give architects and historians new tools for confirming known theories or even discovering new ones. The work was published in [13] and the results are illustrated in figure 7.

7.2.5. Proposal Flow

Participants: Bumsub Ham, Minsu Cho, Cordelia Schmid, Jean Ponce.

Finding image correspondences remains a challenging problem in the presence of intra-class variations and large changes in scene layout, typical in scene flow computation. In [22], we introduce a novel approach to this problem, dubbed proposal flow, that establishes reliable correspondences using object proposals. Unlike prevailing scene flow approaches that operate on pixels or regularly sampled local regions, proposal flow benefits from the characteristics of modern object proposals, that exhibit high repeatability at multiple scales,

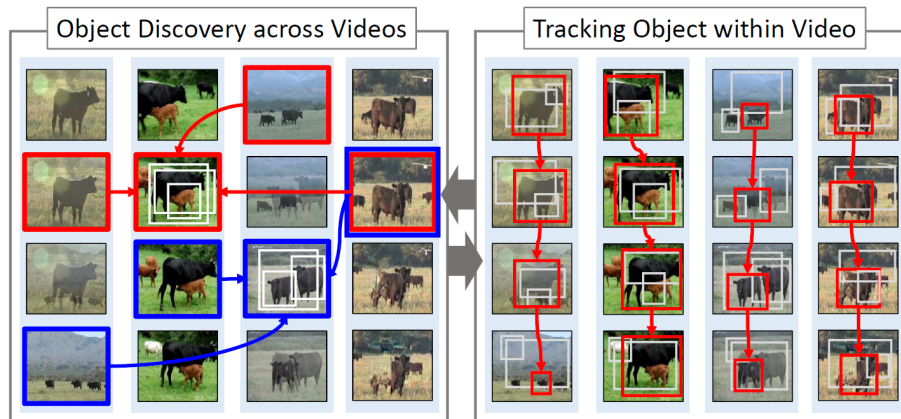


Figure 6. Dominant objects in a video collection are discovered by analyzing correspondences between prominent regions across videos (left). Within each video, object candidates, discovered by the former process, are temporally associated and a smooth spatio-temporal localization is estimated (right). These processes are alternated until convergence or up to a fixed number of iterations.

and can take advantage of both local and geometric consistency constraints among proposals. We also show that proposal flow can effectively be transformed into a conventional dense flow field. We introduce a new dataset that can be used to evaluate both general scene flow techniques and region-based approaches such as proposal flow. We use this benchmark to compare different matching algorithms, object proposals, and region features within proposal flow with the state of the art in scene flow. This comparison, along with experiments on standard datasets, demonstrates that proposal flow significantly outperforms existing scene flow methods in various settings. This work is under review. The proposed method and its qualitative result are illustrated in Figure 8.

7.3. Image restoration, manipulation and enhancement

7.3.1. Learning a Convolutional Neural Network for Non-uniform Motion Blur Removal

Participants: Jian Sun, Wenfei Cao, Zongben Xu, Jean Ponce.

In this work, we address the problem of estimating and removing non-uniform motion blur from a single blurry image. We propose a deep learning approach to predicting the probabilistic distribution of motion blur at the patch level using a convolutional neural network (CNN). We further extend the candidate set of motion kernels predicted by the CNN using carefully designed image rotations. A Markov random field model is then used to infer a dense non-uniform motion blur field enforcing the motion smoothness. Finally the motion blur is removed by a non-uniform deblurring model using patch-level image prior. Experimental evaluations show that our approach can effectively estimate and remove complex non-uniform motion blur that cannot be well achieved by the previous approaches. This work has been published at CVPR 2015[15].

7.3.2. Robust Image Filtering Using Joint Static and Dynamic Guidance

Participants: Bumsub Ham, Minsu Cho, Jean Ponce.

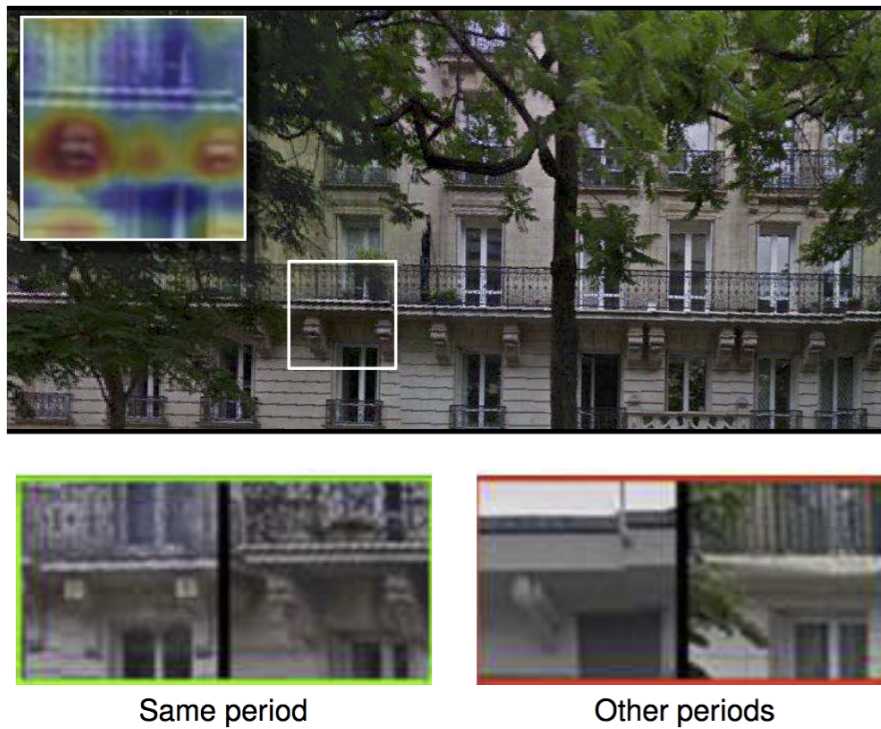


Figure 7. Using thousands of Street View images aligned to a cadastral map, we automatically find visual elements distinctive to particular architectural periods. For example, the patch in white above was found to be distinctive to the Haussmann period (late 1800's) in Paris, while the heat map (inset) reveals that the ornate balcony supports are the most distinctive features. We can also find functionally-similar elements from the same and different time periods (bottom).



Figure 8. Proposal flow generates a reliable scene flow between similar images by establishing geometrically consistent correspondences between object proposals. (Left) Region-based scene flow by matching object proposals. (Right) Color-coded dense flow field generated from the region matches, and image warping using the flow.

Filtering images using a guidance signal, a process called joint or guided image filtering, has been used in various tasks in computer vision and computational photography, particularly for noise reduction and joint upsampling. The aim is to transfer the structure of the guidance signal to an input image, restoring noisy or altered image structure. The main drawbacks of such a data-dependent framework are that it does not consider differences in structure between guidance and input images, and it is not robust to outliers. We propose a novel SD (for static/dynamic) filter to address these problems in a unified framework by jointly leveraging structural information of guidance and input images. Joint image filtering is formulated as a nonconvex optimization problem, which is solved by the majorization-minimization algorithm. The proposed algorithm converges quickly while guaranteeing a local minimum. The SD filter effectively controls the underlying image structure at different scales and can handle a variety of types of data from different sensors. It is robust to outliers and other artifacts such as gradient reversal and global intensity shifting, and has good edge-preserving smoothing properties. We demonstrate the flexibility and effectiveness of the SD filter in a great variety of applications including depth upsampling, scale-space filtering, texture removal, flash/non-flash denoising, and RGB/NIR denoising. This has been published at CVPR 2015 [10]. The SD filter is illustrated in Figure 9 .

7.3.3. PCS-Net: A Deep learning approach to image restoration

Participants: Jian Sun, Jean Ponce.

This work introduces a novel framework for image restoration casting this problem as a joint classification and regression task. This is a learning-based approach, which first classifies degraded image patches into different categories, then restores these patches using category-specific models. We implement this idea by designing a novel convolutional neural network (dubbed PCS-Net), combining a CNN-based patch classification subnet with a novel patch category switched CNN architecture for category-specific restoration. The proposed PCS-Net learns different weights for different patch categories in a common network structure. Experiments on standard benchmarks show that our approach matches or improves upon the state of the art in image super-resolution and denoising. This work is under review.

7.4. Human activity capture and classification

7.4.1. P-CNN: Pose-based CNN Features for Action Recognition

Participants: Guilhem Chéron, Ivan Laptev, Cordelia Schmid.

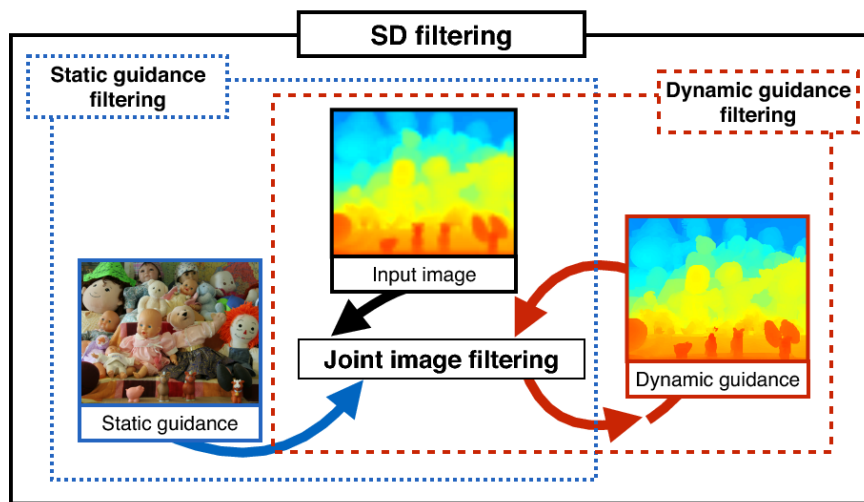


Figure 9. Sketch of joint image filtering and SD filtering: Static guidance filtering convolves an input image with a weight function computed from static guidance, as in the dotted blue box. Dynamic guidance filtering uses weight functions that are repeatedly obtained from regularized input images, as in the dotted red box. We have observed that static and dynamic guidance complement each other, and exploiting only one of them is problematic, especially in the case of data from different sensors (e.g., depth and color images). The SD filter takes advantage of both, and addresses the problems of current joint image filtering.

This work [9] targets human action recognition in video. We argue for the importance of a representation derived from human pose. To this end we propose a new Pose-based Convolutional Neural Network descriptor (P-CNN) for action recognition. The descriptor aggregates motion and appearance information along tracks of human body parts as shown in Figure 10. We experiment with P-CNN features obtained both for automatically estimated and manually annotated human poses. We evaluate our method on JHMDB and MPII Cooking datasets. For both datasets our method shows consistent improvement over the state of the art. This work has been published at ICCV 2015 [9], and P-CNN code (Matlab) is available online at <http://www.di.ens.fr/willow/research/p-cnn/>.

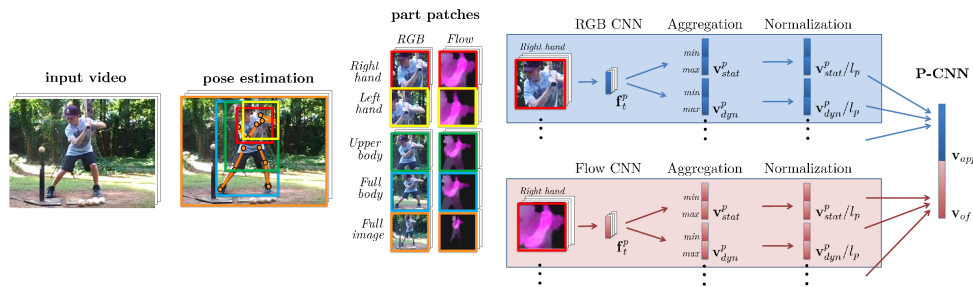


Figure 10. P-CNN features. From left to right: Input video. Human pose. Patches of appearance and optical flow for human body parts. One RGB and one flow CNN descriptor is extracted per frame and per part. Frame descriptors are aggregated over time to obtain the video descriptor. Video descriptors are normalized and concatenated into appearance features and flow features. The final P-CNN feature is the concatenation of appearance and flow.

7.4.2. Context-aware CNNs for person head detection

Participants: Tuan-Hung Vu, Anton Osokin, Ivan Laptev.

Person detection is a key problem for many computer vision tasks. While face detection has reached maturity, detecting people under a full variation of camera view-points, human poses, lighting conditions and occlusions is still a difficult challenge. In this work we focus on detecting human heads in natural scenes. Starting from the recent local R-CNN object detector, we extend it with two types of contextual cues. First, we leverage person-scene relations and propose a Global CNN model trained to predict positions and scales of heads directly from the full image. Second, we explicitly model pairwise relations among objects and train a Pairwise CNN model using a structured-output surrogate loss. The Local, Global and Pairwise models are combined into a joint CNN framework. To train and test our full model, we introduce a large dataset composed of 369,846 human heads annotated in 224,740 movie frames. We evaluate our method and demonstrate improvements of person head detection against several recent baselines in three datasets. We also show improvements of the detection speed provided by our model. This work has been published at ICCV 2015 [18]. The code and the new dataset developed in this work are available online at <http://www.di.ens.fr/willow/research/headddetection/>.

7.4.3. On Pairwise Costs for Network Flow Multi-Object Tracking

Participants: Visesh Chari, Simon Lacoste-Julien, Ivan Laptev, Josef Sivic.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the “tracking-by-detection” paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose an

approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method. This work has been published at CVPR 2015 [7].

7.4.4. Pose Estimation and Segmentation of Multiple People in Stereoscopic Movies

Participants: Guillaume Seguin, Karteek Alahari, Josef Sivic, Ivan Laptev.

We describe a method to obtain a pixel-wise segmentation and pose estimation of multiple people in stereoscopic videos, illustrated in Figure 11. This task involves challenges such as dealing with unconstrained stereoscopic video, non-stationary cameras, and complex indoor and outdoor dynamic scenes with multiple people. We cast the problem as a discrete labelling task involving multiple person labels, devise a suitable cost function, and optimize it efficiently. The contributions of our work are two-fold: First, we develop a segmentation model incorporating person detections and learnt articulated pose segmentation masks, as well as colour, motion, and stereo disparity cues. The model also explicitly represents depth ordering and occlusion. Second, we introduce a stereoscopic dataset with frames extracted from feature-length movies “StreetDance 3D” and “Pina”. The dataset contains 587 annotated human poses, 1158 bounding box annotations and 686 pixel-wise segmentations of people. The dataset is composed of indoor and outdoor scenes depicting multiple people with frequent occlusions. We demonstrate results on our new challenging dataset, as well as on the H2view dataset from (Sheasby et al.’s ACCV 2012). This work has been published at PAMI [4].

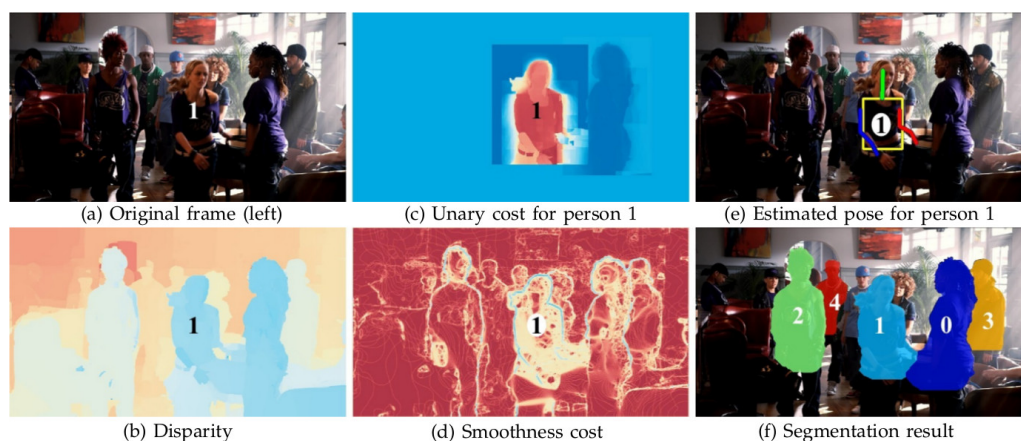


Figure 11. Starting from a stereo pair (a), we estimate disparity maps (b). Using both appearance and disparity cues, we detect persons and estimate their poses (e). We combine pose information with disparity information and occlusion reasoning to compute the unary potentials of a CRF (c) and use standard color and motion cues to compute the binary terms (d). We optimize the CRF problem to produce the final, layered segmentation (f).

7.4.5. Weakly-Supervised Alignment of Video with Text

Participants: Piotr Bojanowski, Rémi Lajugie, Edouard Grave, Francis Bach, Ivan Laptev, Jean Ponce, Cordelia Schmid.

In this work [6], we design a method for aligning natural language sentences with a video stream. Suppose that we are given a set of videos, along with natural language descriptions in the form of multiple sentences (e.g., manual annotations, movie scripts, sport summaries etc.), and that these sentences appear in the same temporal order as their visual counterparts. We propose here a method for aligning the two modalities, i.e., automatically providing a time stamp for every sentence (see Fig. 12). Given vectorial features for both video and text, we propose to cast this task as a temporal assignment problem, with an implicit linear mapping between the two feature modalities. We formulate this problem as an integer quadratic program, and solve its continuous convex relaxation using an efficient conditional gradient algorithm. Several rounding procedures are proposed to construct the final integer solution. After demonstrating significant improvements over the state of the art on the related task of aligning video with symbolic labels, we evaluate our method on a challenging dataset of videos with associated textual descriptions, using both bag-of-words and continuous representations for text. This work has been published at CVPR 2015 [6].

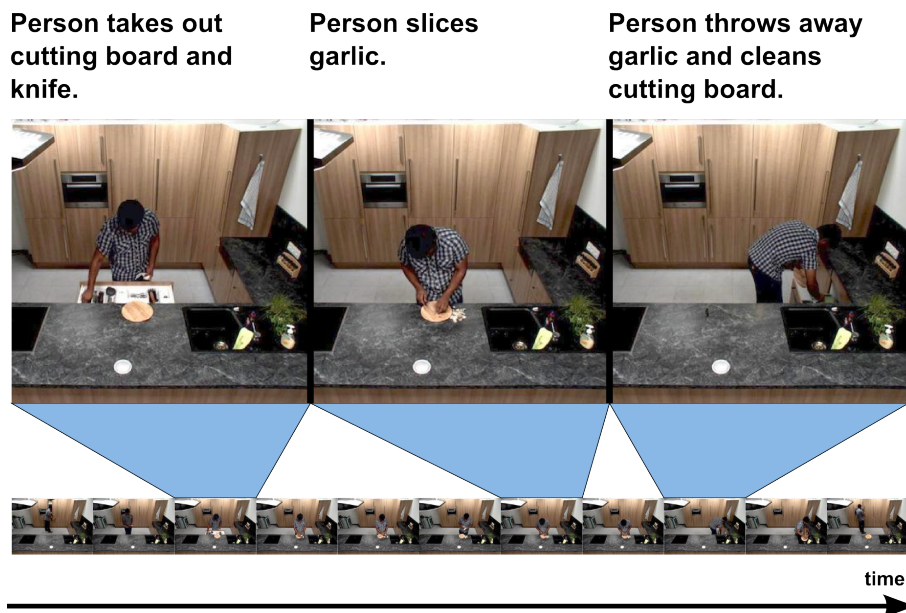


Figure 12. Illustration of the text to video alignment problem. As an output, our model provides a temporal location for every sentence.

7.4.6. Unsupervised learning from narrated instruction videos

Participants: Jean-Baptiste Alayrac, Piotr Bojanowski, Nishant Agrawal, Josef Sivic, Ivan Laptev, Simon Lacoste-Julien.

In [20], we address the problem of automatically learning the main steps to complete a certain task, such as changing a car tire, from a set of narrated instruction videos. The contributions of this paper are three-fold. First, we develop a new unsupervised learning approach that takes advantage of the complementary nature of the input video and the associated narration. The method solves two clustering problems, one in text and one in video, applied one after each other and linked by joint constraints to obtain a single coherent sequence of steps in both modalities. Second, we collect and annotate a new challenging dataset of real-world instruction videos from the Internet. The dataset contains about 800,000 frames for five different tasks that include complex interactions between people and objects, and are captured in a variety of indoor and outdoor settings. Third, we

experimentally demonstrate that the proposed method can automatically discover, in an unsupervised manner, the main steps to achieve the task and locate the steps in the input videos. This work is under review.

7.4.7. Long-term Temporal Convolutions for Action Recognition

Participants: Gül Varol, Ivan Laptev, Cordelia Schmid.

Typical human actions such as hand-shaking and drinking last several seconds and exhibit characteristic spatio-temporal structure. Recent methods attempt to capture this structure and learn action representations with convolutional neural networks. Such representations, however, are typically learned at the level of single frames or short video clips and fail to model actions at their full temporal scale. In [27], we learn video representations using neural networks with long-term temporal convolutions. We demonstrate that CNN models with increased temporal extents improve the accuracy of action recognition despite reduced spatial resolution. We also study the impact of different low-level representations, such as raw values of video pixels and optical flow vector fields and demonstrate the importance of high-quality optical flow estimation for learning accurate action models. We report state-of-the-art results on two challenging benchmarks for human action recognition UCF101 and HMDB51. This work is under review. The results for the proposed method are illustrated in Figure 13 .



Figure 13. The highest improvement of long-term temporal convolutions in terms of class accuracy is for “JavelinThrow”. For 16-frame network, it is mostly confused with “FloorGymnastics” class. We visualize sample videos with 7 frames extracted at every 8 frames. The intuitive explanation is that both classes start by running for a few seconds and then the actual action takes place. Long-term temporal convolutions with 60 frames can capture this interval, whereas 16-frame networks fail to recognize such long-term activities.

7.4.8. Thin-Slicing for Pose: Learning to Understand Pose without Explicit Pose Estimation

Participants: Suha Kwak, Minsu Cho, Ivan Laptev.

In [23], we address the problem of learning a pose-aware, compact embedding that projects images with similar human poses to be placed close-by in the embedding space (Figure 14). The embedding function is built on a deep convolutional network, and trained with a triplet-based rank constraint on real image data. This

architecture allows us to learn a robust representation that captures differences in human poses by effectively factoring out variations in clothing, background, and imaging conditions in the wild. For a variety of pose-related tasks, the proposed pose embedding provides a cost-efficient and natural alternative to explicit pose estimation, circumventing challenges of localizing body joints. We demonstrate the efficacy of the embedding on pose-based image retrieval and action recognition problems. This work is under review.

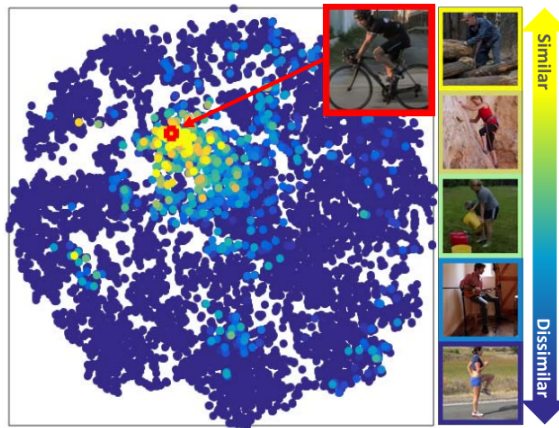


Figure 14. The manifold of our pose embedding visualized using *t*-SNE. Each point represents a human pose image. To better show correlation between the pose embedding and annotated pose, we color-code pose similarities in annotation between an arbitrary target image (red box) and all the other images. Selected examples of color-coded images are illustrated in the right-hand side. Images similar with the target in annotated pose are colored in yellow, otherwise in blue. As can be seen, yellow images lie closer by the target in general, which indicates that a position on the embedding space implicitly represents a human pose.

7.4.9. Instance-level video segmentation from object tracks

Participants: Guillaume Seguin, Piotr Bojanowski, Rémi Lajugie, Ivan Laptev.

In [26], we address the problem of segmenting multiple object instances in complex videos. Our method does not require manual pixel-level annotation for training, and relies instead on readily-available object detectors or visual object tracking only. Given object bounding boxes at input as shown in Figure 15, we cast video segmentation as a weakly-supervised learning problem. Our proposed objective combines (a) a discriminative clustering term for background segmentation, (b) a spectral clustering one for grouping pixels of same object instances, and (c) linear constraints enabling instance-level segmentation. We propose a convex relaxation of this problem and solve it efficiently using the Frank-Wolfe algorithm. We report results and compare our method to several baselines on a new video dataset for multi-instance person segmentation. This work is under review.



Figure 15. Results of our method applied to multi-person segmentation in a sample video from our database. Given an input video together with the tracks of object bounding boxes (left), our method finds pixel-wise segmentation for each object instance across video frames (right).

WIMMICS Project-Team

7. New Results

7.1. Users modeling and designing interaction

7.1.1. Exploratory search

Participants: Emilie Palagi, Alain Giboin.

Contrary to lookup search engines that help users to retrieve specific items (e.g., names, numbers, short statements, or specific documents), Exploratory Search Systems (ESSs) are search engines that help users to explore a topic of interest. Exploratory search (ES) tasks are open-ended, multi-faceted, and iterative like learning or topic investigation [59]. Currently, the evaluation methods of ESSs are not entirely adapted to the special features of ES tasks, and do not effectively assess that ESSs support users in performing those tasks. Our research goal is to elaborate methods that effectively lead to this assessment. Two research actions were undertaken this year to contribute to achieve this goal.

7.1.1.1. Design of an exploratory-search-oriented protocol for testing an image search algorithm based on user's eye movements

Participants: Emilie Palagi, Alain Giboin.

This work was undertaken in the context of the VISIIR ANR project⁰, led by the MinD team (I3S, UNS), with Stéphanie Lopez and Frédéric Precioso. One of the objectives of VISIIR is to design an interactive image search system based on user's eye movements. Detected by an eye-tracker, these movements allow the system to infer the image that the user is going to select; VISIIR's aim is to replace user's mouse clicks as a selection mode by implicit *eye-clicks*. Since ES behaviors can be observed in image search tasks, we designed an eye-tracking user test protocol on Discovery Hub⁰ in order to: 1) increase our understanding of the ES process (at the cognitive and perceptual-motor levels); 2) verify if identified characteristics of gaze trajectories allow to infer the images that the user is going to select in ES tasks (as opposed to lookup search tasks).

7.1.1.2. Design of a user-centered evaluation method of exploratory search systems based on a model of the exploratory search process

Participants: Emilie Palagi, Alain Giboin, Fabien Gandon.

(with Raphaël Troncy, Eurecom)

This work was undertaken in the context of the PhD of Emilie Palagi. In [41] we introduced our approach for designing a user-centered evaluation method for ESSs. Our method takes into account users's ES behavior and is based on a cognitive model of an ES task. We will specially work on Discovery Hub (Wimmics project – Inria) and 3cixty⁰ (EURECOM project) ESSs.

7.1.2. Sentiment Analysis

Participant: Andrea Tettamanzi.

Together with Célia da Costa Pereira (I3S, UNS) and Mauro Dragoni of FBK, Trento, who visited our team for three months from April to June 2014, we have further refined our approach to concept-level sentiment analysis based on fuzzy logic [12].

7.1.3. Recommendation of Pedagogical Resources Adapted to User Profile and Context

Participants: Oscar Rodriguez Rocha, Catherine Faron-Zucker.

⁰<http://www.agence-nationale-recherche.fr/?Projet=ANR-13-CORD-0009>

⁰<http://discoveryhub.co/>

⁰<https://www.3cixty.com/>

In the framework of the Semantic Educloud project, we developed a Web ontology for the description and representation of serious games. Such ontology describes the functional and design elements of the game, the profile and virtual context of the players and furthermore the datasets from the Web of data that the game can query. The ontology has been evaluated through a prototype, which is basically a serious game quiz based on DBpedia. As future work, it is planned to implement state-of-the-art recommendation algorithms of Linked Data resources that take into account the context and players' profile. Furthermore an integration with the EDUCLOUD platform is considered: EDUCLOUD is an emerging initiative in Sophia Antipolis, for the implementation of a platform of digital educational content accessible through the cloud from a 3D portal of resources, and any interface devices (tablets, smartphones, PCs). [44]

7.2. Communities and social interactions analysis

7.2.1. Community Detection and Interest Labeling

Participants: Zide Meng, Fabien Gandon, Catherine Faron-Zucker.

7.2.1.1. Topic Modeling Based Overlapping Community Detection

Based on previous work, we conducted more experiments to evaluate the effectiveness and efficiency of the proposed tag tree based method. We used perplexity score to evaluate the performance of topic extraction. We got consistent performance when applying the model on a Flickr dataset. This work has been published in IEEE/WIC/ACM Web Intelligence 2015 [32] and Social Network Analysis and Mining Journal [13].

7.2.1.2. Temporal Analysis of User and Topic

By jointly modeling topic, expertise, time and activity, we were able to retrieve many meaningful latent information from the user generated contents. We proposed a method to track the dynamics of topics and users. It can also track the dynamics with a specific granularity of time level such as, yearly, monthly, daily and hourly. Besides, the model can overcome a comparison problem of LDA based model by modeling the reverse distribution.

7.2.1.3. Topic labeling

The output of topic model is normally a bag of words. Each topic consists of closely related words. An interesting question is to assign one or more topic label to this set in order to indicate the general meaning of a bag of words. By integrating the original dataset with linked open data sources, we are now planning to propose a generic method to automatically label the detected topics.

7.2.2. Semantic Modeling of Social, Spatiotemporal and Dedicated Networks

Participants: Amel Ben Othmane, Nhan Le Thanh, Michel Buffa, Andrea Tettamanzi, Serena Villata.

We have been working on modeling a multi-agent based recommender system. The aim of such system is to recommend a list of activities (plans) according to user preferences in order to achieve a goal. For this purpose, we propose a multi-context framework based on the well-known agent Belief-Desire-Intention (BDI) architecture [58]. First, we extend the BDI model with additional contexts in order to handle sociality. Second, we use a possibilistic approach based on the work of Da Costa Pereira & Tettamanzi [55], to reason about beliefs, desires, goals and intentions. Further, we use ontologies to represent and reason about plans and intentions. The proposed framework is detailed in a long paper that will be presented in the 8th International Conference on Agents and Artificial Intelligence in 2016 [18].

7.2.3. Collaborative Software Development Platforms

Participant: Isabelle Mirbel.

The collaborative nature of software development helped in the emergence of several online collaborative software development platforms (CSDPs). These platforms enable distributed teams of contributors to participate in the development of the various hosted projects. In such a context, the identification of relevant contributors is very important for handling efficiently the abundant requirements. However, this can be really challenging because of the fairly large number of involved contributors, especially in some distinguished projects. Moreover, the contributor profiles on a CSDP are often inadequately informative, which makes them an unqualified resource for learning about the contributors.

In this context, we proposed to identify contributors by their reputation on a CSDP. Our approach calculates reputation scores using a belief calculus, called subjective logic, according to contributors' performed roles. Knowing the reputation of anonymous contributors would enable project members to reduce the uncertainty in their future interactions with them. Moreover, we use concept lattices to classify contributors by their reputation scores, which enable us to have a comparable view on the considered contributors. Consequently, we can produce a roadmap to examine new requirements thus supporting their effective communication and prioritization.

7.2.4. Logical Foundations of Cognitive Agents

Participants: Andrea Tettamanzi, Serena Villata.

Together with Célia da Costa Pereira (I3S, UNS), we have continued an investigation about the issue of trust in multi-agent systems, and we proposed a computational model of trust based on the content of messages and on the characteristics of their sources [27].

7.2.5. Combining Argumentation Theory and Normative Reasoning with Natural Language Processing

Participants: Serena Villata, Elena Cabrio.

Together with Cristian Cardellino and Laura Alonso Alemany from the University of Cordoba (Argentina), we applied different Active Learning strategies to Information Extraction from licenses in English, with highly repetitive text, few annotated or unannotated examples available, and very fine precision needed. We showed that the most popular approach to active learning, i.e., uncertainty sampling for instance selection, does not provide a good performance in this setting. We showed that we can obtain a similar effect to that of density-based methods using uncertainty sampling, by just reversing the ranking criterion, and choosing the most certain instead of the most uncertain instances. The results of this research have been published at the CICLing [24] and JURIX [23] international conferences.

In another work, together with Alessio Palmero Arosio (FBK Trento, Italy), we have worked on an extension of QAKiS, the system for open domain Question Answering over linked data, that allows to query DBpedia multilingual chapters. Such chapters can contain different information with respect to the english version, e.g. they provide more specificity on certain topics, or fill information gaps. In particular, we have introduced and evaluated the RADAR 2.0 framework for information reconciliation over language-specific DBpedia chapters. The framework is composed of three main modules: a module computing the confidence score of the sources depending either on the length of the related Wikipedia page or on the geographical characterization of the queried entity, a module retrieving the relations holding among the elements of the results set, and finally a module computing the reliability degree of such elements depending on the confidence assigned to the sources and the relations among them. This third module is based on bipolar argumentation theory to return the acceptability degrees. A demo of the RADAR framework is available online⁰. This contribution has been submitted to the Semantic Web Journal and is under review.

Moreover, we have proposed the BEGincs (BEG-Inconsistencies) framework, which translates a bipolar entailment graph into an argumentation graph. It then provides to the annotators sets of arguments that are supposed to be consistent. If it is not the case, the Textual Entailment system wrongly assigned some relations. Moving from single pairs to an overall graph allows for the detection of inconsistencies otherwise undiscovered. BEGincs does not identify the precise relation causing the inconsistency, but provides annotators with the consistent arguments sets, they are supported in narrowing the causes of inconsistency. The results of this research have been published at the CLIC conference [43].

7.2.6. Argumentation theory and its applications

Participants: Elena Cabrio, Serena Villata, Fabien Gandon, Andrea Tettamanzi.

⁰<http://qakis.org/qakis2>

Together with Celia da Costa Pereira (UNS), we have proposed a framework to measure the acceptability of an information in a multiagent system, according to (i) the agent's goals and the information source's goals, (ii) the credibility, for the agent, of the incoming information and (iii) the agent's beliefs (or perceptions) about the context (or situation) in which it operates. The results of this research have been published at the AAMAS international conference [27].

Moreover, together with Sahbi Benlamine, Maher Chaouachi and Claude Frasson (U. of Montreal) we have presented an empirical evaluation of the relationship between the argumentative structures of human debates and the emotions felt by the debate participants. Argumentation is often seen as a mechanism to support different forms of reasoning such that decision-making and persuasion, but all these approaches assume a purely rational behavior of the involved actors. However, humans are proved to behave differently, mixing rational and emotional attitudes to guide their actions, and it has been claimed that there exists a strong connection between the argumentation process and the emotions felt by people involved in such process. We assess this claim by means of an experiment: during several debates people's argumentation in plain English is connected and compared to the emotions automatically detected from the participants. The results of this research have been published at the IJCAI international conference [19], and submitted to the Cognitive Science journal (under review).

7.2.7. *Natural Language Argumentation on Twitter*

Participants: Tom Bosc, Elena Cabrio, Serena Villata.

A great amount of textual data is published on social media every day. For example, there are about 500 million new tweets per day on Twitter. These data reflect the opinion and thoughts of a large population and are thus potentially useful to decision-makers and marketers, among others. But processing them is challenging because of their large quantity as well as their noisiness : poor quality of writing, redundancy, presence of advertisement, etc.

The goal of this project is to build a pipeline to automatically analyze messages exchanged on Twitter and build informative and synthetic views. We study tweets under the angle of argumentation theory. First of all, the algorithm filters in argumentative tweets. Then, it describes how tweets relate to one another : tweets may support or attack other tweets, or be neutral. Finally, a visualisation of the interactions between tweets is produced. Individual parts of the pipeline are machine learning models that are trained using datasets that are crafted specially for the project. Importantly, datasets span several domains (politics, society topics, product announcements) to ensure that the approach is generic enough and will generalize to unseen topics.

7.3. Vocabularies, Semantic Web and linked data based knowledge representation

7.3.1. *SPARQL Template Transformation Language*

Participants: Olivier Corby, Catherine Faron-Zucker, Fabien Gandon, Fuqi Song.

We designed and developed a generic software environment to generate Semantic Web Servers and Linked Data Navigators ⁰ on top of the STTL SPARQL Template Transformation Language. We designed STTL transformations from RDF to HTML that enable to set up hypertext Linked Data Navigators on local or remote (e.g. DBpedia) triple stores. This work was published at ISWC, WebIST, LNBIP and IC [26], [25], [45], [39].

We extended STTL in order to perform rule based constraint checking. Templates return boolean true (resp. false) when constraint checking succeeds (resp. fails). We applied this extension on OWL profile conformance checking and we tested with success OWL RL, OWL EL and OWL QL profiles.

7.3.2. *SPARQL Function Language*

Participants: Olivier Corby, Catherine Faron-Zucker.

⁰<http://corese.inria.fr>

We started the design of a Function Language on top of SPARQL filter language. We added the function statement that enables users to define extension functions directly in the filter language. We added statements to the filter languages such as `let` local variables, `for` loop and `list` datatype and we integrated `select` and `construct` queries in the language. Extension functions are directly available into SPARQL queries. This solves the problem of extension function interoperability. We were able to design custom datatypes such as roman numbers, custom aggregates such as median and standard deviation, extension functions to compute the week day of a given date, approximate search functions, recursive functions with the `service` clause, etc. [50].

7.3.3. Graph Pattern Matching

Participants: Olivier Corby, Fuqi Song.

We proposed a heuristics-based query planning approach which allows reducing SPARQL query executing time. This approach has been developed and integrated to Corese platform. The relevant work and results have been published at conference KES 2015 [35].

We developed a component that can improve the storage capacity of Corese software, generally speaking this approach stores large RDF literals into the file system instead of in memory. The experiments are performed based on the data set of BSBM [54] and the results suggested that with this component, it can save up to 40% RAM space without slowing down the query execution time.

We implemented and integrated similarity measurement algorithms to Corese software in order to enable approximate semantic search. The main objective is to return approximate results when there are no results in the data source corresponding to the query.

7.3.4. Dynamic Application Scheme Composition

Participant: Isabelle Mirbel.

Dynamic service composition has emerged as a promising approach to build complex runtime-adaptable applications. In this context, new approaches for bottom-up opportunistic assembly of services have emerged. However, these approaches may lead to meaningless and useless compositions. Therefore, we advocate an approach in which bottom-up discovery of services is coupled with top-down user's requirements elicitation.

In our approach, application schemes publish available behaviors from basic component assembly. Our user's requirements elicitation framework, based on previous work, offers the capability to capture high-level end-user's requirements in an iterative and incremental way and to turn them into queries to retrieve application scheme's descriptions. We adopt semantic Web languages and models as a unified framework to deal with end-user's requirements and application scheme's descriptions in order to take advantage of their reasoning and traceability capabilities. We extended previous work about requirement's modeling by providing means to represent and reason on AND and OR operators as well as contextual data. Moreover, relying on the STTL language (see Section 7.3.1), we proposed two transformations for runtime composition: the first transformation aims at detecting the possible compositions with regards to the available applications schemes; the second one aims at building a BPMN modeling to achieve user's requirements.

7.3.5. Semantic Web Languages And Techniques for Digital Humanities

Participants: Serena Villata, Elena Cabrio, Catherine Faron-Zucker, Franck Michel.

In the framework of the Zoomathia project, we conducted three complementary works. Their results have been published at the SW4SH international workshop [22][37][38]. First, together with Cécile Callou, Chloé Martin and Johan Montagnat (UNS), we started working on the construction of a thesaurus to support multi-disciplinary studies on the transmission of zoological knowledge throughout historical periods, combining the analysis of ancient literature, iconographic and archaeozoological resources. We constructed a SKOS thesaurus based on the TAXREF taxonomical reference designed to support studies in Conservation Biology.

Second, together with Molka Tounsi (UNS), and Arnaud Zucker (UNS), we have introduced a methodology to (i) extract pertinent knowledge from medieval texts using Natural Language Processing methods, (ii) semantically enrich semi-structured zoological data and publishing it as an RDF dataset and its vocabulary, linked to other relevant Linked Data sources, and (iii) reason on this linked RDF data to help epistemologists, historians and philologists in their analysis of these ancient texts.

Third, together with Arnaud Zucker, we have proposed to adopt argumentation theory together with Semantic Web languages and techniques to provide an overall view of conflicting critiques over ancient texts, and to detect what are the different competing viewpoints and what are the strongest arguments emerging from the debate. An ontology for argumentative documents is used to annotate ancient texts, and an example of such annotation is provided about the topic of the Eternity of the species in Aristotle.

Moreover, together with Ahmed Missaoui (UNS) and Sara Tonelli (FBK Trento, Italy), we have presented the process performed to map the metadata from the Verbo-Visual-Virtual Project to the Linked Open Data cloud and the related data enrichment. Although the work was largely inspired by past efforts by other cultural heritage institutions, they face new challenges, partly related to the small size of the collection, with little-known artists and few information available from other online sources, and partly to the integration of Natural Language Processing techniques to enrich the metadata. The results of this research have been published at the AIUCD international conference.

7.3.6. *Autonomous Learning of the Meaning of Objects*

Participants: Valerio Basile, Elena Cabrio, Fabien Gandon.

The goal of ALOOF (CHIST-ERA) project ⁰ is to enable robots to tap into the ever-growing amount of knowledge available on the Web, by learning from there about the meaning of previously unseen objects, expressed in a form that makes them applicable when acting in situated environments. By searching the Web, robots will be able to learn about new objects, their specific properties, where they might be stored and so forth. To achieve this, robots need a mechanism for translating between the representations used in their real-world experience and those on the Web.

In this direction, we are building a *machine reading* pipeline to extract formally encoded knowledge from unstructured text. By combining linguistic and semantic analysis of natural language with entity linking and formal reasoning techniques, our system is capable of extracting meaningful knowledge about entities with URIs in the Linked Open Data (e.g., from DBpedia) and their relationships, encoded in standard Semantic Web fashion, i.e., RDF triples. We then employ the machine reading software to harvest the Web, targeting informative natural language resources such as educational Websites, to create a large-scale meaning bank of common sense knowledge.

7.3.7. *Social Media Intelligence and Linked Knowledge*

Participants: Farhad Nooralahzadeh, Elena Cabrio, Fabien Gandon.

Automated Natural Language Processing (NLP), Web Open Data (Linked Open Data) and social networks are the three topics of the SMILK ANR LabCom including their coupling studied in three ways: texts and Linked Data, Linked Data and social resources, texts and social resources. It is a Joint laboratory between the Inria research institute and the VISEO company to develop research and technologies on the one hand, retrieve, analyze, and reason about linking data from textual Web resources and other to use open Web data taking into account the social structures and interactions in order to improve the analysis and understanding of textual resources.

In this context, we have developed the entity discovery tools by adopting the semantic spreading activation, then we integrated it with the SMILK framework. The goal of this work was to semantically enrich the data by linking the mentions of named entities in the text to the corresponding known entities in knowledge bases. In our approach multiple aspects are considered: the prior knowledge of an entity in Wikipedia (i.e. the keyphraseness and commonness features that can be precomputed by crawling the Wikipedia dump), a set of

⁰<http://www.dis.uniroma1.it/~alooof/>

features extracted from the input text and from the knowledge base, along with the correlation/relevancy among the resources in Linked Data. More precisely, this work explores the *collective ranking approach* formalized as a weighted graph model, in which the mentions in the input text and the candidate entities from knowledge bases are linked using the local compatibility and the global relatedness. Experiments on the datasets of the Open Knowledge Extraction (OKE)⁰ challenge with different configurations of our approach in each phase of the linking pipeline reveal its optimum mode. We investigate the notion of semantic relatedness between two entities represented as sets of neighbors in Linked Open Data that relies on an associative retrieval algorithm, with consideration of common neighborhood. This measure improves the performance of prior link-based models and outperforms the explicit inter-link relevancy measure among entities (mostly Wikipedia-centric). Thus, our approach is resilient to non-existent or sparse links among related entities.

7.3.8. *Ontology-Based Workflow Management Systems*

Participants: Tuan Anh Pham, Nhan Le Thanh.

The main objective of this PhD work is to develop a Shared Workflow Management System (SWMS) using ontology engineering. Everybody can share a semi-complete workflow which is called “Workflow template”, and other people can modify and complete it to use it in their system. This customized workflow is called “Personalized workflow”. The challenges of a SWMS are to be simple, easy to use, user friendly and not too heavy. But it must have all functions of a WMS. There are three major challenges in this work: How to allow the users to customize the workflow template to correspond to their requirements, with changes compliant with the predefined rules in the workflow template? How to build an execution model to evaluate step by step a personalized workflow [34][33].

7.3.9. *Semantic Mappings with a Control Flow-Based Business Workflow*

Participants: Thi Hoa Hue Nguyen, Nhan Le Thanh.

The aim of this PhD work is to improve the Coloured Petri Nets (CPNs) and Ontology engineering to support the development of business process and business workflow definitions of various fields. To realize this objective, we first propose an ontological approach for representing business models in a meta-knowledge base. We introduce four basic types of manipulation operations on process models used to develop and modify business workflow patterns. Second, we propose a formal definition of semantic constraints and an $O(n^3)$ -time algorithm for detecting redundant and conflicting constraints. By relying on the CPN Ontology and sets of semantic constraints, workflow processes are semantically created. Finally, we show how to check the semantic correctness of workflow processes with the SPARQL query language [34].

7.4. Analyzing and Reasoning on Heterogeneous Semantic Graphs

7.4.1. *RDF Mining*

Participants: Andrea Tettamanzi, Catherine Faron-Zucker, Fabien Gandon, Tran Duc Minh, Claudia d’Amato.

We carried on our investigation in an approach to RDF mining based on grammatical evolution and possibility theory, whose aim is to mine large RDF graphs by automatically generating and testing OWL 2 axioms based on the known facts. In particular, we addressed the problem of scaling up the scoring heuristics based on falsification and possibility theory we have recently proposed [36].

7.4.2. *Data and Knowledge Integration and Extraction*

Participant: Andrea Tettamanzi.

Together with Somsack Inthasone of the National University of Laos, Nicolas Pasquier and Célia da Costa Pereira of I3S, we completed a survey on biodiversity and environment data mining [16].

⁰<https://github.com/anuzzolese/oke-challenge>

7.4.3. Scalable Uncertainty Management

Participant: Andrea Tettamanzi.

Within the framework of the CNR PEPS GéoIncertitude, we proposed and studied the properties of uncertain logical gates in possibilistic network, using a problem of human geography as a motivating example and testbed [28].

7.4.4. Natural Language Question Answering

Participants: Andrea Tettamanzi, Elena Cabrio, Catherine Faron-Zucker, Amine Hallili.

We extended previous work on answering N -relation natural language questions in the commercial domain by combining an approach to learning regular expressions based on genetic programming [21].

7.4.5. Events Detection in Twitter

Participants: Amosse Edouard, Elena Cabrio, Nhan Le Thanh.

We analyze Twitter data in the objective of identifying events reported by Twitter users. Specially we have worked on two main aspects: an approach for classifying tweets as either related or not related to events and secondly we have studied an approach for disambiguating geographic entities in tweets.

We have worked on an approach for separating event-related content from the rest of Twitter messages. We have combined technics from Natural Language Processing (NLP) and Machine Learning (ML) for building a classifier model that aims at classifying tweets into two mutually exclusive classes. First of all, we apply a Named Entity Recognizer to the tweets in order to identify the occurrences of named entities and special Twitter features such as hashtags, shortened URLs or user mentions. In a second step, the named entities are replaced by their generic class in the DBpedia Ontology; we do so by using SPARQL to query the DBpedia Knowledge Base to extract the class related to each entity. Third, we use the modified content as examples to train a binary classifier. Our evaluation using different classifiers such as Naive Bayes and Long Short Term Memory have shown promising results in term of performance compared to the state of the art.

We have also worked on an approach for identifying geographic entities in Twitter. This task is challenging for two main reasons: first, a geographic term can be related to either geographic or non geographic entities (Paris can be a person or a place) and second, many geographic places might have the same name (Paris can be either the capital of France or a city in Texas). We have proposed an approach based on distant-supervision and ontology matching for identifying and disambiguate ambiguous geographic terms.

ZENITH Project-Team

7. New Results

7.1. Big Data Integration

7.1.1. *CloudMdsQL, a query language for heterogeneous data stores*

Participants: Carlyna Bondiombouy, Boyan Kolev, Oleksandra Levchenko, Patrick Valduriez.

The blooming of different cloud data management infrastructures, specialized for different kinds of data and tasks, has led to a wide diversification of DBMS interfaces and the loss of a common programming paradigm. The CoherentPaaS European project addresses this problem, by providing a common programming language and holistic coherence across different cloud data stores.

In this context, we have started the design of a Cloud Multi-datastore Query Language (CloudMdsQL), and its query engine. CloudMdsQL is a functional SQL-like language, capable of querying multiple heterogeneous data stores, e.g. relational, NoSQL or HDFS) [19], [31]. The major innovation is that a CloudMdsQL query can exploit the full power of the local data stores, by simply allowing some local data store native queries to be called as functions, and at the same time be optimized. Our experimental validation, with three data stores (graph, document and relational) and representative queries, shows that CloudMdsQL satisfies the five important requirements for a cloud multidatastore query language. In [32], we extend CloudMdsQL to allowing the ad-hoc usage of user defined map/filter/reduce operators in combination with traditional SQL statements, to integrate relational data and big data stored in HDFS and accessed by a data processing framework like Spark.

7.1.2. *Semantic Data Integration using Bio-Ontologies*

Participant: Pierre Larmande.

The AgroPortal project [49] aims at developing and supporting a reference ontology repository for the agronomic domain. The ontology portal features ontology hosting, search, versioning, visualization, comment, with services for semantically annotating data with the ontologies, as well as storing and exploiting ontology alignments and data annotations. All of these within a fully semantic web compliant infrastructure. The main objective of this project is to enable straightforward use of agronomic related ontologies, avoiding data managers and researchers the burden to deal with complex knowledge engineering issues to annotate the research data. Thus, we specifically pay attention to the requirements of the agronomic community and the specificities of the crop domain. AgroPortal will offer a robust and stable platform that we anticipate will be highly valued by the community.

7.1.3. *Access and Integration of Molecular Biology Data*

Participants: Sarah Cohen-Boulakia, Patrick Valduriez.

The volumes of molecular biology data available on the web are constantly increasing. Accessing and integrating these data is crucial for making progress in biology. In [26], we provide all the necessary pointers to identify the reference databases capable of providing bioinformatic data for molecular biology. We also discuss the problems posed by the exploitation of these very highly heterogeneous and distributed data. Finally, in order to guide a prospective user on the choice of one of these systems, we provide an overview of the systems that provide unified access to these data.

7.2. Distributed Indexing and Searching

7.2.1. *Diversified and Distributed Recommendation for Scientific Data*

Participants: Esther Pacitti, Maximilien Servajean.

Recommendation is becoming a popular mechanism to help users find relevant information in large-scale data (scientific data, web). To avoid redundancy in the results, recommendation diversification has been proposed, with the objective of identifying items that are dissimilar, but nonetheless relevant to the user's interests.

We propose a new diversified search and recommendation solution suited for scientific data (i.e., plant phenotyping, botanical data) [22]. We first define an original profile diversification scoring function that enables to address the problem of returning redundant items, and enhances the quality of diversification. Through experimental evaluation using two benchmarks, we showed that our scoring function gives the best compromise between diversity and relevancy. Next, to implement our new scoring function, we propose a basic Top-k threshold-based algorithm that exploits a candidate list to achieve diversification and several techniques to improve performance. First, we simplify the scoring model to reduce its computational complexity. Second, we propose two techniques to reduce the number of items in the candidate list, and thus the number of diversified scores to compute. Third, we propose different indexing scores that take into account the diversification of items and an adaptive indexing approach to reduce the number of accesses in the index dynamically based on the queries workload. The experimentation results show that our techniques yield a major reduction of response time, up to 12 times compared to a baseline greedy diversification algorithm.

We also address distributed and diversified recommendation in the context of P2P and multisite cloud [23]. We propose a new scoring function (usefulness) to cluster relevant users over a distributed overlay. Our experimental evaluation using different datasets shows major gains in recall (order of 3 times) compared with state-of-the-art solutions.

7.3. Scientific Workflows

7.3.1. Scientific Workflows: combining data analysis and simulation

Participant: Sarah Cohen-Boulakia.

While scientific workflows are increasingly popular in the bioinformatics community in some emerging application domains such as ecology, the need for data analysis is combined with the need to model complex multi-scale biological systems, possibly involving multiple simulation steps. This requires the scientific workflow to deal with retro-action to understand and predict the relationships between structure and function of these complex systems. OpenAlea (openalea.gforge.inria.fr) developed by the EPI Virtual plants is the only scientific workflow system able to uniformly address the problem, which made it successful in the scientific community.

For the first time, we proposed a conceptualisation of OpenAlea in [42]. We introduce the concept of higher-order dataflows as a means to uniformly combine classical data analysis with modeling and simulation. We provide for the first time the description of the OpenAlea system involving an original combination of features. We illustrate the demonstration on a high-throughput workflow in phenotyping, phenomics, and environmental control designed to study the interplay between plant architecture and climatic change. Ongoing work include deploying OpenAlea on a Grid technology using the SciFloware middleware.

7.3.2. Processing Scientific Workflows in Multi-site cloud

Participants: Ji Liu, Esther Pacitti, Patrick Valduriez.

As the scale of the data increases, scientific workflow management systems (SWfMSs) need to support workflow execution in High Performance Computing (HPC) environments. Because of various benefits, cloud emerges as an appropriate infrastructure for workflow execution. However, it is difficult to execute some scientific workflows in one cloud site because of geographical distribution of scientists, data and computing resources. Therefore, a scientific workflow often needs to be partitioned and executed in a multisite environment.

In [21], we define a multisite cloud architecture that is composed of traditional clouds, e.g., a pay-per-use cloud service such as Amazon EC2, private data-centers, e.g. a cloud of a scientific organization like Inria, COPPE or LNCC, and client desktop machines that have authorized access to the data-centers. We can model this architecture as a distributed system on the Internet, each site having its own computer cluster, data and programs. An important requirement is to provide distribution transparency for advanced services (i.e., workflow management, data analysis), to ease their scalability and elasticity. Current solutions for multisite clouds typically rely on application specific overlays that map the output of one task at a site to the input of another in a pipeline fashion. Instead, we define fully distributed services for data storage, intersite data movement and task scheduling.

7.3.3. *Data-centric Iteration in Dynamic Workflows*

Participant: Patrick Valduriez.

Dynamic workflows are scientific workflows supporting computational science simulations, typically using dynamic processes based on runtime scientific data analyses. They require the ability of adapting the workflow, at runtime, based on user input and dynamic steering. Supporting data-centric iteration is an important step towards dynamic workflows because user interaction with workflows is iterative. However, current support for iteration in scientific workflows is static and does not allow for changing data at runtime.

In [17], we propose a solution based on algebraic operators and a dynamic execution model to enable workflow adaptation based on user input and dynamic steering. We introduce the concept of iteration lineage that makes provenance data management consistent with dynamic iterative workflow changes. Lineage enables scientists to interact with workflow data and configuration at runtime through an API that triggers steering. We evaluate our approach using a novel and real large-scale workflow for uncertainty quantification on a 640-core cluster. The results show impressive execution time savings from 2.5 to 24 days, compared to non-iterative workflow execution. We verify that the maximum overhead introduced by our iterative model is less than 5% of execution time. Also, our proposed steering algorithms are very efficient and run in less than 1 millisecond, in the worst-case scenario.

7.3.4. *Analyzing Related Raw Data Files through Dataflows*

Participant: Patrick Valduriez.

Computer simulations may ingest and generate high numbers of raw data files. Most of these files follow a de facto standard format established by the application domain, e.g., FITS for astronomy. Although these formats are supported by a variety of programming languages, libraries and programs, analyzing thousands or millions of files requires developing specific programs. DBMS are not suited for this, because they require loading the raw data and structuring it, which gets heavy at large-scale. Systems like NoDB, RAW and FastBit, have been proposed to index and query raw data files without the overhead of using a DBMS. However, they focus on analyzing one single large file instead of several related files. In this case, when related files are produced and required for analysis, the relationship among elements within file contents must be managed manually, with specific programs to access raw data. Thus, this data management may be time-consuming and error-prone. When computer simulations are managed by a SWfMS, they can take advantage of provenance data to relate and analyze raw data files produced during workflow execution. However, SWfMS register provenance at a coarse grain, with limited analysis on elements from raw data files. When the SWfMS is dataflow-aware, it can register provenance data and the relationships among elements of raw data files altogether in a database which is useful to access the contents of a large number of files. In [24], we propose a dataflow approach for analyzing element data from several related raw data files. Our approach is complementary to the existing single raw data file analysis approaches. We validate our approach with the Montage workflow from astronomy and a workflow from Oil and Gas domain as I/O intensive case studies.

7.4. Scalable Query Processing

7.4.1. *Scalable Query Processing with Big Data*

Participants: Reza Akbarinia, Miguel Liroz, Patrick Valduriez.

The popular MapReduce parallel processing framework is inefficient in case of data skew, which makes the reduce side done by a few worker nodes.

In [28], [20], we propose FP-Hadoop, which makes the reduce side of MapReduce more parallel. We extend the MapReduce programming model to allow the collaboration of reduce workers on processing the values of an intermediate key, without affecting the correctness of the final results. In FP-Hadoop, the reduce function is replaced by two functions: intermediate reduce and final reduce. There are three phases, each phase corresponding to one of the functions: map, intermediate reduce and final reduce phases. In the intermediate reduce phase, the function, which usually includes the main load of reducing in MapReduce jobs, is executed by reduce workers in a collaborative way, even if all values belong to only one intermediate key. This allows performing a big part of the reducing work by using the computing resources of all workers, even in case of highly skewed data. We implemented a prototype of FP-Hadoop by modifying Hadoop's code, and conducted extensive experiments over synthetic and real datasets. The results show that FP-Hadoop makes MapReduce job processing much faster and more parallel, and can efficiently deal with skewed data. We achieve excellent performance gains compared to native Hadoop, e.g. more than 10 times in reduce time and 5 times in total execution time.

7.5. Data Stream Mining

7.5.1. Summarizing Uncertain Data Streams

Participants: Reza Akbarinia, Florent Masegla.

Probabilistic data management has shown growing interest to deal with uncertain data. In [29], we focus on probabilistic time series with high volumes of data, which requires efficient compression techniques. To date, most of the work on probabilistic data reduction uses synopses that minimize the error of representation wrt. the original data. However, in most cases, the compressed data will be meaningless for usual queries involving aggregation operators such as SUM or AVG. We propose *PHA* (Probabilistic Histogram Aggregation), a compression technique whose objective is to minimize the error of such queries over compressed probabilistic data. We incorporate the aggregation operator given by the end-user directly in the compression technique, and obtain much lower error in the long term. We also adopt a global error aware strategy in order to manage large sets of probabilistic time series, where the available memory is carefully balanced between the series, according to their individual variability.

7.6. Scalable Data Analysis

7.6.1. Parallel Mining of Maximally Informative k -Itemsets in Big Data

Participants: Saber Salah, Reza Akbarinia, Florent Masegla.

The discovery of informative itemsets is a fundamental building block in data analytics and information retrieval. While the problem has been widely studied, only few solutions scale. This is particularly the case when i) the data set is massive, and/or ii) the length K of the informative itemset to be discovered is high. In [45], we address the problem of parallel mining of maximally informative k -itemsets (miki) based on joint entropy. We propose PHIKS (Parallel Highly Informative K -itemSets) a highly scalable, parallel mining algorithm. PHIKS renders the mining process of large scale databases (up to terabytes of data) succinct and effective. Its mining process is made up of only two compact, yet efficient parallel jobs. PHIKS uses a clever heuristic approach to efficiently estimates the joint entropies of miki having different sizes with very low upper bound error rate, which dramatically reduces the runtime process. PHIKS has been extensively evaluated using massive, real-world data sets. Our experimental results confirm the effectiveness of our approach by the significant scale-up obtained with high featuresets length and hundreds of millions of objects.

7.6.2. Frequent Itemset Mining in Massively Distributed Environments

Participants: Saber Salah, Reza Akbarinia, Florent Masegla.

While the problem of Frequent itemset mining (FIM) has been thoroughly studied, few solutions scale. This is mainly the case when i) the amount of data tends to be very large and/or ii) the minimum support (MinSup) threshold is very low. In [46], we study the effectiveness and leverage specific data placement strategies for improving parallel FIM (PFIM) performance in MapReduce, a highly distributed computation framework. By offering a clever data placement and an optimal organization of the extraction algorithms, we show that the itemset discovery effectiveness does not only depend on the deployed algorithms. We propose ODPR (Optimal Data-Process Relationship), a solution for fast mining of frequent itemsets in MapReduce. Our method allows discovering itemsets from massive datasets, where standard solutions do not scale.

In [44], we propose a highly scalable PFIM algorithm, namely Parallel Absolute Top Down (PATD). PATD renders the mining process of very large databases (up to Terabytes) simple and compact. Its mining process is made up of only one parallel job, which dramatically reduces the mining runtime, communication cost and energy power consumption overhead, in a distributed computational platform. Based on a clever and efficient data partitioning strategy, namely Item Based Data Partitioning (IBDP), PATD mines each data partition independently, relying on an absolute minimum support (AM inSup) instead of a relative one. Through an extensive experimental evaluation using real-world data sets, we show that PATD is significantly more efficient and scalable than alternative approaches.

7.6.3. Scalable Mining of Closed Frequent Itemsets

Participants: Mehdi Zitouni, Reza Akbarinia, Florent Masseglia.

Mining big datasets poses a number of challenges which are not easily addressed by traditional mining methods, since both memory and computational requirements are hard to satisfy. One solution is to take advantage of parallel frameworks, such as MapReduce, using ordinary machines. In [48], we address the issue of mining closed frequent itemsets (CFI) from big datasets in such environments. We introduce a new parallel algorithm, called CloPN, for CFI mining. One important feature of CloPN is to use a prime number based approach to transform the data into numerical form, and then to mine closed frequent itemsets by using only multiplication and division operations. We carried out exhaustive experiments over big real world datasets to assess the performance of CloPN. The results show that our algorithm is very efficient in CFI mining from large real world datasets with up to 53 million articles.

7.6.4. Chiaroscuro

Participants: Tristan Allard, Florent Masseglia, Esther Pacitti.

The advent of on-body/at-home sensors connected to personal devices leads to the generation of fine grain highly sensitive personal data at an unprecedented rate. However, despite the promises of large scale analytics there are obvious privacy concerns that prevent individuals to share their personal data. In [30], we propose Chiaroscuro, a complete solution for clustering personal data with strong privacy guarantees. The execution sequence produced by Chiaroscuro is massively distributed on personal devices, coping with arbitrary connections and disconnections. Chiaroscuro builds on our novel data structure, called Diptych, which allows the participating devices to collaborate privately by combining encryption with differential privacy. Our solution yields a high clustering quality while minimizing the impact of the differentially private perturbation. Our study shows that Chiaroscuro is both correct and secure.

7.6.5. Large-scale Recognition of Visual and Audio Entities

Participants: Valentin Leveau, Alexis Joly, Patrick Valduriez.

We improved our work on the retrieval of visual identities by introducing a supervised classification layer on top of the large-scale instance-based matching layer. We introduce a new match kernel based on the inverse rank of the Shared Nearest Neighbors (SNN) combined with local geometric constraints [40]. To avoid overfitting and reduce processing costs, the dimensionality of the resulting over-complete representation is further reduced by hierarchically pooling the raw consistent matches according to their spatial position in the training images. The final image representation is obtained by concatenating the resulting feature vectors at several resolutions. Learning from these representations using a logistic regression classifier is shown to

provide excellent fine-grained classification performance. In [38], we transpose our new SNN match kernel to the case of audio contents (applied to bird sounds recognition). Thus, the spatial pooling of geometrically consistent visual matches is replaced by a temporal pooling of temporally consistent audio matches. The resulting classification system obtained the second best results at the LifeCLEF bird identification challenge 2015 [36], the largest challenge of this kind ever organized (1000 bird species, 33K audio recordings).

7.6.6. Crowd-sourced Biodiversity Data Production through PI@ntNet

Participants: Alexis Joly, Julien Champ, Jean-Christophe Lombardo, Antoine Affouard.

Initiated in the context of a citizen sciences project with botanists of the AMAP laboratory and the Tela Botanica social network, PI@ntNet [18] is an innovative collaborative platform focused on image-based plant identification as a mean to enlist new contributors and boost the production of biodiversity data and knowledge. Since 2010, several hundreds of thousands of geo-tagged and dated plant photographs were collected and revised by tens of thousands of novice, amateur and expert botanists. A content-based identification tool, available as both web and mobile applications, is synchronized with the growing data and allows any user to query or enrich the system with new observations. As a concrete new result, the cumulative number of downloads of the iPhone or Android app did reach 1M in October 2015. One of the main novelty in 2015 was the introduction of deep learning technologies in order to improve classification performance as well as the quality and speed of the content-based image retrieval.

A comparative study that we conducted in the context of the LifeCLEF⁰ plant identification challenge did actually confirm that deep convolutional neural networks definitely outperforms the best fine-grained classification models on the aggregation of hand-crafted visual features [33]. Thus, we integrated this technology in the PI@ntNet platform and exploited it in two ways: (i) for extracting more relevant (local and global) visual features to be indexed and searched within our efficient content-based indexing and retrieval framework (SnoopIm software) (ii) for reranking the species returned by the content-based search engine so as to increase the average reciprocal rank of the correct species while keeping a good level of interpretability of the returned results.

7.6.7. Crowd-sourced Biodiversity Data Production through LifeCLE

Participants: Alexis Joly, Julien Champ, Jean-Christophe Lombardo, Antoine Affouard.

We continued sharing the data produced by the PI@ntNet platform with the international research community through the animation of the LifeCLEF research platform and the set-up of three new challenges, one related to plant images, one to bird sounds and one to fish videos. More than 200 research groups registered to at least one of the challenges and about 20 of them crossed the finish lines by running their system on the final test data. A synthesis of the results is published in the LifeCLEF 2015 overview paper [37] and more detailed analyses are provided in technical reports for the plant task [35] and the bird task [36]. We also report on an experimental study aimed at evaluating how state-of-art computer vision systems perform in identifying plants compared to human expertise [15]. A subset of the evaluation dataset used within LifeCLEF 2014 plant identification challenge was shared with volunteers of diverse expertise, ranging from leading experts of the targeted flora to inexperienced test subjects. In total, 16 human runs were collected and evaluated comparatively to the 27 machine-based runs of LifeCLEF challenge. The main outcome of the experiment was that machines are still far from outperforming the best expert botanists but they are clearly competing with some experienced botanists specialists of other floras.

⁰www.lifeclef.org