RESEARCH CENTER
Paris - Rocquencourt

FIELD

Activity Report 2013

Section New Results

Edition: 2014-03-19

1. ABSTRACTION Project-Team 4
2. ALPAGE Project-Team 10
3. ALPINES Team 19
4. ANGE Team 23
5. AOSTE Project-Team 26
6. ARAMIS Team 32
7. ARLES Project-Team 47
8. AXIS Project-Team 54
9. BANG Project-Team 67
10. CAD Team 75
11. CASCADE Project-Team (section vide) 77
12. CLASSIC Project-Team 78
13. CLIME Project-Team 80
14. CONTRAINTES Project-Team 90
15. CRYPT Team (section vide) 95
16. DEDUCTEAM Exploratory Action 96
17. DYOGENE Project-Team 100
18. FORMES Team 111
19. GALLIUM Project-Team 113
20. GAMMA3 Project-Team 123
21. GANG Project-Team 127
22. HIPERCOM2 Team 135
23. IMARA Project-Team 141
24. MATHRISK Project-Team 149
25. MICMAC Project-Team 153
26. MOKAPLAN Exploratory Action 159
27. MUTANT Project-Team 165
28. PARKAS Project-Team 169
29. PI.R2 Project-Team 175
30. POLSYS Project-Team 180
31. POMDAPI Project-Team (section vide) 187
32. PROSECCO Project-Team 188
33. RAP Project-Team 191
34. REGAL Project-Team 196
35. REO Project-Team 202
36. SECRET Project-Team 209
37. SIERRA Project-Team 214
38. SISYPHE Project-Team 227
39. SMIS Project-Team 232
40. WILLOW Project-Team 234

ABSTRACTION Project-Team

6. New Results

6.1. Analysis of Biological Pathways

We have improved our framework to design and analyze biological networks in KAPPA. This framework focuses on protein-protein interaction networks described as graph rewriting systems. Such networks can be used to model some signaling pathways that control the cell cycle. The task is made difficult due to the combinatorial blow up in the number of reachable species (*i.e.*, non-isomorphic connected components of proteins).

6.1.1. Semantics

Participants: Jonathan Hayman, Tobias Heindel [CEA-List].

Keywords: Graph rewriting, Single Push-Out semantics.

Domain-specific rule-based languages can be understood intuitively as transforming graph-like structures, but due to their expressivity these are difficult to model in 'traditional' graph rewriting frameworks.

In [16], we introduce pattern graphs and closed morphisms as a more abstract graph-like model and show how Kappa can be encoded in them by connecting its single-pushout semantics to that for Kappa. This level of abstraction elucidates the earlier single-pushout result for Kappa, teasing apart the proof and guiding the way to richer languages, for example the introduction of compartments within cells.

6.1.2. Causality Analysis

We use causal analysis so as to extract minimal concurrent scenarios that lead to the activation of given events.

6.1.2.1. Implementation

Participant: Jérôme Feret.

Keywords: Causality, Counter-examples, Compression.

This year, we have re-implemented in OPENKAPPA the strong compression method that is described in [48]. The new implementation is very efficient, it has been used to extract minimal scenarios from traces of several hundred of thousands causally related events, that were generated during the simulation of a model of the WnT signaling pathway.

6.1.2.2. Framework

Participant: Jonathan Hayman.

Keywords: Abstraction, Causality, Compression.

Standard notions of independence of rule applications fail to provide adequately concise causal histories, leading to the earlier formulation of strong and weak forms of trajectory compression for Kappa. In [15], we give a simple categorical account of how forms of compression can be uniformly obtained. This generalisation also describes a way for the user to specify their own levels of compression between weak and strong, which we call filtered compression. This is based on the idea of the user specifying the part of the type graph that represents the the structure which the compression technique should track through the trace.

6.1.3. Model Reduction

Participants: Ferdinanda Camporesi, Jérôme Feret, Jonathan Hayman.

Keywords: Context-sensitivity, Differential semantics, Model reduction.

Rule-based modeling allows very compact descriptions of protein-protein interaction networks. However, combinatorial complexity increases again when one attempts to describe formally the behaviour of the networks, which motivates the use of abstractions to make these models more coarse-grained. Context-insensitive abstractions of the intrinsic flow of information among the sites of chemical complexes through the rules have been proposed to infer sound coarse-graining, providing an efficient way to find macro-variables and the corresponding reduced models.

In [12], we propose a framework to allow the tuning of the context-sensitivity of the information flow analyses and show how these finer analyses can be used to find fewer macro-variables and smaller reduced differential models.

6.2. Andromeda: Accurate and Scalable Security Analysis of Web Applications

Participants: Omer Tripp [Tel Aviv University, Israël], Marco Pistola [University of Washington, Seattle, USA], Patrick Cousot, Radhia Cousot, Salvatore Guarnieri.

Keywords: Abstract interpretation, Security, Web.

Security auditing of industry-scale software systems mandates automation. Static taint analysis enables deep and exhaustive tracking of suspicious data flows for detection of potential leakage and integrity violations, such as cross-site scripting (XSS), SQL injection (SQLi) and log forging. Research in this area has taken two directions: program slicing and type systems. Both of these approaches suffer from a high rate of false findings, which limits the usability of analysis tools based on these techniques. Attempts to reduce the number of false findings have resulted in analyses that are either (i) unsound, suffering from the dual problem of false negatives, or (ii) too expensive due to their high precision, thereby failing to scale to real-world applications.

In [21], we investigate a novel approach for enabling precise yet scalable static taint analysis. The key observation informing our approach is that taint analysis is a demand-driven problem, which enables lazy computation of vulnerable information flows, instead of eagerly computing a complete data-flow solution, which is the reason for the traditional dichotomy between scalability and precision. We have implemented our approach in Andromeda, an analysis tool that computes data-flow propagations on demand, in an efficient and accurate manner, and additionally features incremental analysis capabilities. Andromeda is currently in use in a commercial product. It supports applications written in Java, .NET and JavaScript. Our extensive evaluation of Andromeda on a suite of 16 production-level benchmarks shows Andromeda to achieve high accuracy and compare favorably to a state-of-the-art tool that trades soundness for precision.

6.3. Backward analysis

6.3.1. Automatic Inference of Necessary Preconditions

Participants: Patrick Cousot, Radhia Cousot, Manuel Fähndrich [Microsoft Research, Redmond, USA], Francesco Logozzo [Microsoft Research, Redmond, USA].

Keywords: Abstract interpretation, Backward analysis, Static analysis, Necessary condition inference.

In [14], we consider the problem of automatic precondition inference for: (i) program verification; (ii) helping the annotation process of legacy code; and (iii) helping generating code contracts during code refactoring. We argue that the common notion of sufficient precondition inference (i.e., under which precondition is the program correct?) imposes too large a burden on call-sites, and hence is unfit for automatic program analysis. Therefore, we define the problem of necessary precondition inference (i.e., under which precondition, if violated, will the program always be incorrect?). We designed and implemented several new abstract interpretation-based analyses to infer necessary preconditions. The analyses infer atomic preconditions (including disjunctions), as well as universally and existentially quantified preconditions.

We experimentally validated the analyses on large scale industrial code.

For unannotated code, the inference algorithms find necessary preconditions for almost 64% of methods which contained warnings. In 27% of these cases the inferred preconditions were also sufficient, meaning all warnings within the method body disappeared. For annotated code, the inference algorithms find necessary preconditions for over 68% of methods with warnings. In almost 50% of these cases the preconditions were also sufficient. Overall, the precision improvement obtained by precondition inference (counted as the additional number of methods with no warnings) ranged between 9% and 21%.

6.3.2. *Under-approximations to infer sufficient program conditions*

Participant: Antoine Miné.

Keywords: Abstract interpretation, Backward analysis, Numerical abstract domains, Static analysis, Sufficient condition inference, Under-approximations.

In [9] we discuss the automatic inference of sufficient preconditions by abstract interpretation and sketch the construction of an under-approximating backward analysis. We focus on numeric properties of variables and revisit three classic numeric abstract domains: intervals, octagons, and polyhedra, with new under-approximating backward transfer functions, including the support for non-deterministic expressions, as well as lower widenings to handle loops. We show that effective under-approximation is possible natively in these domains without necessarily resorting to disjunctive completion nor domain complementation. Applications include the derivation of sufficient conditions for a program to never step outside an envelope of safe states, or dually to force it to eventually fail. We built a proof-of-concept prototype implementation based on the APRON numeric domain library and experimented it on simple examples (the prototype is available for download and usable on-line at http://www.di.ens.fr/~mine/banal).

6.4. Bisimulation metrics

6.4.1. *Bisimulation for MDP through Families of Functional Expressions*

Participants: Norman Ferns, Sophia Knight [LIX], Doina Precup [McGill University].

Keywords: Markov decision processes, Bisimulation, Metrics.

We have transfered a notion of quantitative bisimilarity for labelled Markov processes [54] to Markov decision processes with continuous state spaces. This notion takes the form of a pseudometric on the system states, cast in terms of the equivalence of a family of functional expressions evaluated on those states and interpreted as a real-valued modal logic. Our proof amounts to a slight modification of previous techniques [61], [60] used to prove equivalence with a fixed-point pseudometric on the state-space of a labelled Markov process and making heavy use of the Kantorovich probability metric. Indeed, we again demonstrate equivalence with a fixed-point pseudometric defined on Markov decision processes [57]; what is novel is that we recast this proof in terms of integral probability metrics [59] defined through the family of functional expressions, shifting emphasis back to properties of such families. The hope is that a judicious choice of family might lead to something more computationally tractable than bisimilarity whilst maintaining its pleasing theoretical guarantees. Moreover, we use a trick from descriptive set theory to extend our results to MDPs with bounded measurable reward functions, dropping a previous continuity constraint on rewards and Markov kernels.

This work is under submission.

6.4.2. *Bisimulation Metrics are Optimal Value Functions*

Participants: Norman Ferns, Doina Precup [McGill University].

Keywords: Markov decision processes, Bisimulation, Metrics.

We have proved that a behavioural pseudometric defined on the state space of a given Markov decision process and whose kernel is stochastic bisimilarity [57] can be expressed as the optimal value function of another Markov decision process. Furthermore, this latter process can be interpreted as an optimal coupling of two copies of the original model.

This work is under submission.

6.5. A Constraint Solver Based on Abstract Domains

Participants: Marie Pelleau [University of Nantes, LINA], Antoine Miné, Charlotte Truchet [University of Nantes, LINA], Frédéric Benhamou [University of Nantes, LINA].

Keywords: Abstract interpretation, Backward analysis, Numerical abstract domains, Static analysis, Sufficient condition inference, Under-approximations.

In [18] and [19] we apply techniques from abstract interpretation to constraint programming (which aims at solving hard combinatorial problems with a generic framework based on first-order logics). We highlight some links and differences between these fields: both compute fixpoints by iterations but employ different extrapolation and refinement strategies; moreover, consistencies in constraint programming can be mapped to non-relational abstract domains. We then use these correspondences to build an abstract constraint solver that leverages abstract interpretation techniques (such as relational domains) to go beyond classic solvers. We present encouraging experimental results obtained with our prototype implementation.

6.6. A Galois Connection Calculus for Abstract Interpretation

Participants: Patrick Cousot, Radhia Cousot.

Keywords: Abstract interpretation, Galois connection.

In [10], we introduce a Galois connection calculus for language independent specification of abstract interpretations used in programming language semantics, formal verification, and static analysis. This Galois connection calculus and its type system are typed by abstract interpretation.

6.7. Mechanically Verifying a Shape Analysis

Participant: Arnaud Spiwack.

Keywords: Program verification, Abstract interpretation, Static analysis, Shape analysis, Coq.

The result of a static analysis is only as good as the trust put into its correctness. For critical software, the standards are very high, and trusting a complex tool requires costly inspection of its implementation. Mechanically proving the correctness of static analysers is a way to lower these costs: the exigence of trust is moved from various complex dedicated tools to a single simpler general purpose one.

In this context, Arnaud Spiwack worked on an ongoing Coq implementation and certification of a shape abstract domain. The implementation, named Cosa, is based on Evan Chang and Xavier Rival's Xisa. It targets an intermediary language of Xavier Leroy's Compcert C, and interfaces with the domains of the Verasco project.

The development of Cosa lead Arnaud Spiwack to express the abstract interpretation correctness property in term of refinement calculus, which allowed to use interaction structures (a type theoretic variant of the refinement calculus) as a central structuring element of Cosa. Arnaud Spiwack started investigating how the technology of nominal sets could be leveraged to prove the correctness of unfolding (which involves choosing new names) in Cosa.

6.8. Modular Construction of Shape-Numeric Analyzers

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

Keywords: Abstract interpretation, Memory abstraction, Shape abstract domains.

In [13], we discuss the modular construction of memory abstract domains.

The aim of static analysis is to infer invariants about programs that are precise enough to establish semantic properties, such as the absence of run-time errors. Broadly speaking, there are two major branches of static analysis for imperative programs. Pointer and *shape* analyses focus on inferring properties of pointers, dynamically-allocated memory, and recursive data structures, while *numeric* analyses seek to derive invariants on numeric values. Although simultaneous inference of shape-numeric invariants is often needed, this case is especially challenging and is not particularly well explored. Notably, simultaneous shape-numeric inference raises complex issues in the design of the static analyzer itself.

In this paper, we study the construction of such shape-numeric, static analyzers. We set up an abstract interpretation framework that allows us to reason about simultaneous shape-numeric properties by combining shape and numeric abstractions into a modular, expressive abstract domain. Such a modular structure is highly desirable to make its formalization and implementation easier to do and get correct. To achieve this, we choose a concrete semantics that can be abstracted step-by-step, while preserving a high level of expressiveness. The structure of abstract operations (i.e., transfer, join, and comparison) follows the structure of this semantics. The advantage of this construction is to divide the analyzer in modules and functors that implement abstractions of distinct features.

6.9. Reduced Product Combination of Abstract Domains for Shapes

Participants: Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival, Antoine Toubhans.

Keywords: Abstract interpretation, Memory abstraction, Shape abstract domains.

In [20], we discuss the construction of shape abstract domains by reduced product.

Real-world data structures are often enhanced with additional pointers capturing alternative paths through a basic inductive skeleton (e.g., back pointers, head pointers). From the static analysis point of view, we must obtain several interlocking shape invariants. At the same time, it is well understood in abstract interpretation design that supporting a separation of concerns is critically important to designing powerful static analyses. Such a separation of concerns is often obtained via a reduced product on a case-by-case basis. In this paper, we lift this idea to abstract domains for shape analyses, introducing a domain combination operator for memory abstractions. As an example, we present *simultaneous separating shape graphs*, a product construction that combines instances of separation logic-based shape domains. The key enabler for this construction is a static analysis on inductive data structure definitions to derive relations between the skeleton and the alternative paths. From the engineering standpoint, this construction allows each component to reason independently about different aspects of the data structure invariant and then separately exchange information via a reduction operator. From the usability standpoint, we enable describing a data structure invariant in terms of several inductive definitions that hold simultaneously.

6.10. Relational Thread-Modular Static Value Analysis

Participant: Antoine Miné.

Keywords: Abstract interpretation, Concurrency, Embedded software, Rely-guarantee methods, Run-time errors, Safety.

We study in [17] thread-modular static analysis by abstract interpretation to infer the values of variables in concurrent programs. We show how to go beyond the state of the art and increase an analysis precision by adding the ability to infer some relational and history-sensitive properties of thread interferences. The fundamental basis of this work is the formalization by abstract interpretation of a rely-guarantee concrete semantics which is thread-modular, constructive, and complete for safety properties. We then show that previous analyses based on non-relational interferences can be retrieved as coarse computable abstractions of this semantics; additionally, we present novel abstraction examples exploiting our ability to reason more precisely about interferences, including domains to infer relational lock invariants and the monotonicity of counters. Our method and domains have been implemented in the ASTRÉEA static analyzer (5.3) that checks for run-time errors in embedded concurrent C programs, where they enabled a significant reduction of the number of false alarms.

6.11. Static Analyzers on the Cloud

Participants: Michael Barnett [Microsoft Research, Redmond, USA], Mehdi Bouaziz, Francesco Logozzo [Microsoft Research, Redmond, USA], Manuel Fähndrich [Microsoft Research, Redmond, USA].

A cloud-based static analyzer runs as service. Clients issue analysis requests through the local network or over the internet. The analysis takes advantage of the large computation resources offered by the cloud: the underlying infrastructure ensures scaling and unlimited storage. Cloud-based analyzers may relax performance-precision trade-offs usually associated with desktop-based analyzers. More cores enable more precise and responsive analyses. More storage enables perfect caching of the analysis results, shareable among different clients, and queryable off-line. To realize these advantages, cloud-based analyzers need to be architected differently than desktop ones. In [11], we describe our ongoing effort of moving a desktop analyzer, Clousot, into a cloud-based one, Cloudot.

6.12. Termination

We have explored the analysis of program termination and the inference of sufficient conditions to ensure the definite termination of programs using abstract interpretation techniques. Following [40], we employ a backward analysis over an abstract domain of ranking functions.

6.12.1. Abstract Domain of Segmented Ranking Functions

Participant: Caterina Urban.

We present in [24] and [23] a parameterized abstract domain that infers sufficient conditions for program termination by automatically synthesizing piecewise-defined ranking functions over natural numbers. The analysis uses over-approximations but we prove its soundness, meaning that all program executions respecting these sufficient conditions are indeed terminating. The abstract domain is parameterized by a numerical abstract domain for environments and a numerical abstract domain for functions. This parameterization allows to easily tune the trade-off between precision and cost of the analysis. We describe an instantiation of this generic domain with intervals and affine functions. We define all abstract operators, including widening to ensure convergence. To experiment with this domain, we have implemented a research prototype static analyzer FUNCTION (5.6) that yielded interesting preliminary results.

6.12.2. Abstract Domain to Infer Ordinal-Valued Ranking Functions

Participants: Caterina Urban, Antoine Miné.

We observed that, in some important cases (such as programs with unbounded non-determinism), there does not exist any ranking function over natural numbers. In [22] and [29] we propose a new abstract domain to automatically infer ranking functions over ordinals. We extended the domain of piecewise-defined natural-valued ranking functions introduced in the previous section to polynomials in ω, where the polynomial coefficients are natural-valued functions of the program variables. The abstract domain is parametric in the choice of the maximum degree of the polynomial, and the types of functions used as polynomial coefficients. We have enriched the FUNCTION prototype analyzer (5.6) with an instantiation of our domain using affine functions as polynomial coefficients. We successfully analyzed small but intricate examples that are out of the reach of existing methods. To our knowledge this is the first abstract domain able to reason about ordinals. Handling ordinals leads to a powerful approach for proving termination of imperative programs, which in particular subsumes existing techniques based on lexicographic ranking functions.

ALPAGE Project-Team

6. New Results

6.1. Unsupervised segmentation of Mandarin Chinese

Participants: Pierre Magistry, Benoît Sagot.

In Chinese script, very few symbols can be considered as word boundary markers. The only easily identifiable boundaries are sentence beginnings and endings, as well as positions before and after punctuation marks. Although the script doesn't rely on typography to define (orthographic) "words", a word-level segmentation is often re- quired for further natural language processing, which is a highly non-trivial task.

A great variety of methods have been proposed in the literature, mostly in supervised machine learning settings. Our work addresses the question of unsupervised segmentation, i.e., without any manually segmented training data. Although supervised learning typically performs better than unsupervised learning, we believe that unsuper- vised systems are worth investigating as they require less human labour and are likely to be more easily adaptable to various genres, domains and time periods. They can also provide more valuable insight for linguistic studies.

Amongst the unsupervised segmentation systems described in the literature, two paradigms are often used: Branching Entropy (BE) and Minimum Description Length (MDL). The system we have developed relies on both. We have introduced a new algorithm [22] which searches in a larger hypothesis space using the MDL criterion, thus leading to lower Description Lengths than other previously published systems. Still, this improvement concerning the Description Length does not come with better results on the Chinese word segmentation task, which raises interesting issues. However, it turns out that it is possible to add very simple constraints to our algorithm in order to adapt it to the specificities of Mandarin Chinese in a way that leads to results better than the state-of-the-art on the Chinese word segmentation task.

Moreover, an important part of discrepancies between the various segmentation guidelines concerns the so-called "factoids." This term covers a variety of language phenomena that include: numbers, dates, addresses, email addresses, proper names, and others. We have shown that specific treatment of a subset of such expressions is both sound (as factoids to not resort to general language, which we try and capture with our segmentation model, both rather to conventions that are easy to encode as rules). By augmenting the local grammars of SxPipe to deal with the aforementioned expressionsin Chinese, and use them as a pre-processing for our task, we can discard the matched expressions from the training data and segment them accordingly to the guidelines as a post-processing step. Our results show a significant improvement over previous results.

6.2. Dynamic extension of a French lexical resources based a text stream

Participants: Damien Nouvel, Benoît Sagot, Rosa Stern, Virginie Mouilleron, Marion Baranes.

Lexical incompleteness is a recurring problem when dealing with natural language and its variability. It seems indeed necessary today to regularly validate and extend lexica used by tools processing large amounts of textual data. This is even more true when processing real-time text flows. In this context, we have introduced two series of techniques for addressing words unknown to lexical resources, and applied them to French within the context of the EDyLex ANR project:

- **Extending a morphological lexicon** We have studied neology (from a theoretic and corpus-based point of view) and developed modules for detecting neologisms in AFP news wires in real time and inferring information about them (lemma, category, inflectional class) [24]. We have shown that we are able, using among others modules for analyzing derived and compound neologisms, to generate lexical entries candidates in real time and with a good precision, to be added in the Le*fff* lexicon.
- **Extending an entity database** We have also extended our previous work on named entities detection and linking in order to be able to extract new named entities from AFP news wires and create candidate entries for the *Aleda* entity database.

6.3. Transferring lexical knowledge from a resourced language to a closely-related resource-free language

Participants: Yves Scherrer, Benoît Sagot.

We have developed a generic approach for the transfer of part-of-speech (POS) annotations from a resourced language (RL) towards an etymologically closely related non-resourced language (NRL), without using any bilingual (i.e., parallel) data. We rely on two hypotheses. First, on the lexical level, the two languages share a lot of cognates, i.e., word pairs that are formally similar and that are translations of each other. Second, on the structural level, we admit that the word order of both languages is similar, and that the set of POS tags is identical. Thus, we suppose that the POS tag of one word can be transferred to its translational equivalent in the other language.

The proposed approach consists of two main steps. In the first step, we induce a translation lexicon from monolingual corpora. This step relies on several methods, including a character-based statistical machine translation model to infer cognate pairs, and 3-gram and 4-gram contexts to infer additional word pairs on the basis of their contextual similarity. This step yields a list of $< w_{\mathrm{NRL}}, w_{\mathrm{RL}} >$ pairs. In the second step, the RL lexicon entries are annotated with POS tags with the help of an existing resource, and these annotations are transferred onto the corresponding NRL lexicon entries. We complete the resulting tag dictionary with heuristics based on suffix analogy. This results in a list of $< w_{\mathrm{NRL}}, t >$ pairs, covering the whole NRL corpus.

We have evaluated our methods on several language pairs. We have worked among others onfive language pairs of the Iberic peninsula, where Spanish and Portuguese play the role of RLs: Aragonese–Spanish, Asturian–Spanish, Catalan–Spanish, Galician– Spanish and Galician–Portuguese [27]. We have also conducted experiments on germanic [28] and slavic languages. We have also applied it in a slightly different context, in collaboration with Tomaž Erjavec (IJS, Slovenia), namely that of inducing resources for historical Slovene based on existing resources for contemporary Slovene [26]. Although no direct comparison can be performed, because of the novelty of the task, our results are very satisfying in so far that they are almost as high as published result on a related but simpler task, that of unsupervized part-of-speech tagging — which, contrarily to our work, relies on an existing morphological lexicon for the language at hand.

6.4. Building a large-scale translation graph

Participants: Valérie Hanoka, Benoît Sagot.

Large-scale general-purpose multilingual translation databases are useful in a wide range of Natural Languages Processing (NLP) tasks. This is especially true concerning researches tackling problems specific to under-resourced languages, as translation databases can be used for adapting existing resources in other languages. This has been applied for example for the development of wordnets in languages other than English. There is thus a real need in NLP for *open-source* multilingual lexical databases that compiles as many translations as can be found on any freely available resource in any language.

We have developed, and are about to release, a new open-source heavily multilingual (over 590 languages) translation database built using several sources, namely various wiktionaries and the OPUS parallel corpora.

Our graph was built in several steps. We first extracted a preliminary set of translation and synonym pairs, which we stored in a large translation and synonym graph. We then applied filtering techniques for increasing the accuracy of this graph. We have evaluated the accuracy of our graph as being as high as 98% for translations extracted from wiktionaries.

6.5. Computational morphology

Participant: Benoît Sagot.

In 2013, following previous collaborative work [92], [105], we have designed and developed Alexina$_{\text{PARSLI}}$ in collaboration with Géraldine Walther (LLF and DDL), a formalism for encoding inflectional descriptions (lexicon and grammar) that aims at filling the gap between morphologically and typologically motivated approaches on the one hand and implemented approaches on the other hand, as will be discussed in the remainder of this section. Indeed, Alexina$_{\text{PARSLI}}$ is both:

- an **implementation formalism for PARSLI**, a formal model of inflectional morphology [106] that accounts for concepts underlying the canonical approach of morphological typology;
- an **extension of the Alexina lexical framework** developed at Alpage for modeling lexical information and developing lexical resources. The Alexina framework now supports both morphological grammars that use the original Alexina morphological formalism as well as new grammars developed in Alexina$_{\text{PARSLI}}$.

The Alexina$_{\text{PARSLI}}$ formalism and tools have been proven greatly beneficial to works both in descriptive and formal morphology, in particular in studies about Latin passivisation and Maltese verbal inflection [106] and in studies comparing the compacity of morphological descriptions [106], [92], [105], as well as in NLP, for the efficient development of a large-scale and linguistically sound morphological lexicon for German (a paper describing this new lexicon is to be presented at the LREC 2014 conference).

In collaboration with Géraldine Walther and Guillaume Jacques (CRLAO, CNRS), within operation LR4.11 from strand 6 of the LabEx EFL, we have also developed two Alexina$_{\text{PARSLI}}$ descriptions of (part of) the Khaling (Kiranti, Sino-Tibetan) verbal inflectional system, together with a medium-scale lexicon. Our study shows that an explicit account for the so-called direct-inverse marking, based on concepts developed within PARSLI, allows for a more compact account of this inflectional system [42].

6.6. Extracting Derivational Relations from an Inflectional Lexicon

Participants: Marion Baranes, Benoît Sagot.

Derivational morphological can provide useful information for natural language processing tasks. Indeed, it can improve any application which has to deal with unknown words such as information extraction, spell-checking and others.

We define a morphological family as a set of semantically related lexical entries which differ by their prefix and/or suffix, thus limiting ourselves to concatenative derivational morphology. We shall denote as derivationally related two morphological lexical entries that belong to the same morphological family.

We have developed a system which performs an analogy-based unsupervized extraction of weighted transformation rules that relate derivationally related lexical entries, and use these rules for extracting derivational relations within an existing inflectional lexicon. Our transformation rules can also be used to infer morphological information (both inflectional and derivational) for wordforms unknown to the inflectional lexicon. Our system is language-independent, although restricted to concatenative derivational morphology. We have evaluated it on four languages: English, French, German and Spanish. Our results will be published at the LREC 2014 conference.

6.7. Improving post-OCR correction with shallow linguistic processing

Participants: Kata Gábor, Benoît Sagot.

Providing wider access to national cultural heritage by massive digitalization confronts the actors of the field to a set of new challenges. State of the art optical character recognition (OCR) software currently achieve an error rate of around 1 to 10% depending on the age and the layout of the text. While this quality may be adequate for indexing, documents inteded for reading need to meet higher standards. A reduction of the error rate by a factor of 10 to 100 becomes necessary for the diffusion of digitalized books and journals through emerging technologies such as e-books.

Within the PACTE project, an "Investissements d'avenir" project led by the Numen company, we have worked on the automatic post-processing of digitalized documents in the aim of reducing the OCR error rate by using contextual information and linguistic processing, by and large absent from current OCR engines. At the current stage of the project, we are focusing on French texts coming from the archives of the French National Library (Bibliothèque Nationale de France).

We adopted a hybrid approach, making use of both statistical classification techniques and linguistically motivated modules to detect OCR errors and generate correction candidates. The technology is based on the noisy chanel model, widely used in the field of machine translation and spelling correction and subsequently in OCR post-correction. As to linguistically enhanced models, POS tagging was succesfully applied to spelling correction. However, to our knowledge, little work has been done to exploit linguistic analysis for post-OCR correction.

We have proposed to integrate a shallow processing module to detect certain types of named entities, and a POS tagger trained specifically to deal with NE-tagged input. Our studies demonstrate that linguistically informed processing can efficiently contribute to reduce the error rate by 1) detecting false corrections proposed by the statistical correction module, 2) detecting a certain amount of OCR errors not detected by the statistical correction module.

6.8. Named Entity Linking

Participants: Rosa Stern, Benoît Sagot.

The Ph.D. research work started in 2009 lead in 2013 to the development of a joint entity recognition and linking system for the processing of textual data at the Agence France Presse (AFP).

This system, Nomos, allows to use any existing named entity recognition system, as well as combinations of such systems; their results are passed to a linking module in charge of the association between each detected mention and a unique reference within an existing data inventory. The two tasks (recognition and linking) are jointly operated: the recognition module presents a set of possible detections, which are further disambiguated by the linking module concurrently to the search for the best linking solution to each mention. This joint approach is justified by the need to limit the error propagation between two such modules in a pipeline system.

Experiments were achieved in order to evaluate the performance of Nomos over AFP news wires. They showed that the joint approach, relatively to a purely sequential one, improves the system's global precision, i.e. the linking accuracy as well as the named entity recognition task itself. A gain of 3 points (87,6) is observed for the recognition precision with a low recall loss, while a gain of 8 points (92,9) is observed when several recognition systems are combined - although with a more significant loss of recall.

The Nomos system also allows to anchor of the AFP's textual production in the Linked Data network and the Semantic Web paradigm, since the annotations derived from the entity linking associate each entity to an identified resource in repositories such as Wikipedia, DBPedia, Geonames or the New York Times Linked Data.

6.9. Treebanking at Alpage

Participants: Djamé Seddah, Benoît Sagot, Marie-Hélène Candito, Corentin Ribeyre, Benoît Crabbé, Éric Villemonte de La Clergerie, Virginie Mouilleron, Vanessa Combet.

Since the advents of supervized methods for building accurate statistical parsing models, treebank engineering has become of crucial importance. In fact building a treebank, namely a set of carefully annotated syntactic parses with possibly different annotation layers and covering potentially different text domains, can be seen as providing a parser with both a grammar and a set of probabilities used for disambiguation. The main problem of such approaches lies in the nature of the lexical probabilities: they force the parsing model to be extremely sensitive to its training data and hence limit its performance to some low upper-bound when applied in out-of-domain scenario.

6.9.1. Written French Treebanks

Originating from the merging of two NLP teams specialized in grammar engineering and in which the creation of the first treebank for French was initiated [46], it is no wonder that we decided to increase the coverage of our French Treebank-based parsers by building out-of-domain treebanks: the Sequoia Corpus, [55], [18], made from Europarl, biomedical and wikipedia data, and the French Social Media Bank (outside English, the first data set covering Facebook, Twitter and other social media noisy text data) [95], [96]. We built those two corpus for two purposes: first, we wanted to evaluate the performance of our nlp chains (tokenization, tagging, parsing) on out-of-domain data, being noisy or not ; then we increased the coverage of our French treebank based models by simply adding those new data set to the canonical training set (using of-course many lexical variation, morphological clustering, brown clustering, etc.). We're also on the process of finalizing a new 2600 sentence data set, made essentially of questions, which are strikingly absent from all the treebanks we've been using and developing. So far, only one such data set exist and only for English: the Question-Bank [66]. Our very preliminary results show that simply adding a third of that corpus to the French Treebank greatly improve our parser performance.

Finally, Alpage is leading, in collaboration with the Nancy-based team Calligrame, a project to annotate the Sequoia corpus and the French Treebank with a richer, "deeper" syntactic layer, at the interface between syntax and semantics. A paper describing this effort is to appear at the LREC 2014 conference.

6.9.2. Spoken French Treebank

In collaboration with Anne Abeillé (LLF, CNRS), we have also contributed to the deign of a spoken treebank for French based on data produced in the ANR ETAPE. Contrary to other languages such as English, where spoken treebanks such as the Switchboard corpus treebank (Meteer, 1995), there is no sizable spoken corpus for French annotated for syntactic constituents and grammatical functions. Our project is to build such a resource which will be a natural extension of the Paris 7 treebank (Abeillé et al. 2003) for written French, in order to be able to compare with similar annotations written and spoken French. We have reused and adapted the parser (Petrov et al., 2006) which has been trained on the written treebank, with manual correction and validation. The first results are promising [32].

6.10. Linear time constituent parser

Participant: Benoît Crabbé.

We have designed an efficient and accurate lexicalized LR inspired discriminative parsing algorithm that recasts some current advances in dependency parsing to the constituency setting. We specifically designed and evaluated a Graph Structured Stack-based parser (Huang et al. 2010) using some additional specific approximate inference techniques such as the max violation update for the perceptron (Huang et al. 2012) . By contrast with dependency parsing however, lexicalized constituent parsing raises some additional correctness issues that motivate the explicit use of an LR automata instead of a simpler shift reduce framework.

The parsing model is linear in time and has been evaluated on French data, where it turns out to be state of the art on SPMRL 2013 datasets [29] both in time and in accuracy. The parsing framework has been designed to be further extended with compositional semantic representations and allows in principle an easy integration of ressources — such as those developped in the team — considered to be important for parsing morphologically rich languages.

6.11. Improving FRMG through partially supervised learning

Participant: Éric Villemonte de La Clergerie.

Since the emergence of several statistical parsers for French developed on the French TreeBank (FTB), including those developed at Alpage, it was important to be able to compare the symbolic meta-grammar-based parser FRMG with these statistical parsers on their native treebank, but also possibly to extend the comparison for other treebanks.

A first necessary step in this direction was a conversion from FRMG's native dependency scheme into FTB's dependency scheme, a tedious task highlighting the differences in design at all levels (segmentation, parts of speech, representation of the syntactic phenomena, etc.). A preliminary evaluation has shown that accuracy is good, but largely below the scores reached by the statistical parsers.

A challenge was then to explore if training on the FTB could be used to improve the accuracy of a symbolic parser like FRMG. However, the main difficulty arises from the fact that FTB's dependency scheme has little in common with FRMG's underlying grammar, and that no reverse conversion from FTB to FRMG structures is available. Such a conversion could be investigated but would surely be difficult to develop. Instead, we tried to exploit directly FTB data, using only very minimal assumptions, nevertheless leading to important gains and results close to those obtained by the statistical parsers [31]: it was possible to tune the disambiguisation process of FRMG and strongly increase its accuracy, from 83% up to 87.17% (in terms of CONLL Labeled Attachment Score), a level comparable to those reached by statistical parsers trained on the FTB. Preliminary experiments show that (a) disambiguisation tuning also improve the performances on other corpora and (b) that FRMG seems to be more stable than statistical parsers on corpora other than the FTB. Finer-grained comparison of FRMG wrt statistical parsers have been done that provide some insight for further improvements of FRMG.

The interest is that the technique should be easily adaptable for training data with different annotation schemes. Furthermore, our motivation was not just to improve the performances on the FTB and for the annotation scheme of FTB, for instance by training a reranker (as often done for domain adaptation), but to exploit the FTB to achieve global improvement over all kinds of corpora and for FRMG native annotation scheme.

6.12. Statistical parsing of Morphologically Rich Languages

Participants: Djamé Seddah, Marie-Hélène Candito, Éric Villemonte de La Clergerie, Benoît Sagot.

6.12.1. The SPMRL shared task

Since several years, Djamé Seddah, together with Marie-Hélène Candito and more generally the whole Alpage team, has played a major role in setting up and animating an international network of researchers focusing on parsing morphologically rich languages (MRLs).

In 2013, Djamé Seddah led the organization of the first shared task on parsing MRLs, hosted by the fourth SPMRL workshop and described in a 36-page overview paper that constitutes an in-depth state-of-the-art analysis and review of the domain [29]. The primary goal of this shared task was to bring forward work on parsing morphologically ambiguous input in both dependency and constituency parsing, and to show the state of the art for MRLs. We compiled data for as many as 9 languages, which represents an immense scientific and technical challenge.

6.12.2. DyALog-SR

The SPMRL 2013 shared task was the opportunity to develop and test, with promising results, a simple beam-based shift-reduce dependency parser on top of the tabular logic programming system DYALOG. We used (Huang and Sagae, 2010) as the starting point for this work, in particular using the same simple arc-standard strategy for building projective dependency trees. The parser was also extended to handle ambiguous word lattices, with almost no loss w.r.t. disambiguated input, thanks to specific training, use of oracle segmentation, and large beams. We believe that this result is an interesting new one for shift-reduce parsing.

The current implementation scales correctly w.r.t. sentence length and, to a lesser extent, beam size. Nevertheless, for efficiency reasons, we plan to implement a simple C module for beam management to avoid the manipulation in DYALOG of sorted lists. Interestingly, such a module, plus the already implemented model manager, should also be usable to speed up the disambiguation process of DYALOG-based TAG parser FRMG (de La Clergerie, 2005a). Actually, these components could be integrated in a slow but on-going effort to add first-class probabilities (or weights) in DYALOG, following the ideas of (Eisner and Filardo, 2011) or (Sato, 2008).

6.12.3. The Alpage-LIGM French parser

The second Alpage system that participated to the SPMRL shared task, although on French language only, was developed in collaboration with Mathieu Constant (LIGM), based on the Bonsai architecture. This system is made of several single statistical dependency parsing systems whose outputs are combined into a reparser. We use two types of single parsing architecture: (a) pipeline systems; (b)"joint" systems.

The pipeline systems first perform multi-word expression (MWE) analysis before parsing. The MWE analyzer merges recognized MWEs into single tokens and the parser is then applied on the sentences with this new tokenization. The parsing model is learned on a gold training set where all marked MWEs have been merged into single tokens. For evaluation, the merged MWEs appearing in the resulting parses are expanded, so that the tokens are exactly the same in gold and predicted parses.

The "joint" systems directly output dependency trees whose structure comply with the French dataset annotation scheme. Such trees contain not only syntactic dependencies, but also the grouping of tokens into MWEs, since the first component of an MWE bears dependencies to the subsequent components of the MWE with a specific label. At that stage, the only missing information is the POS of the MWEs, which we predict by applying a MWE tagger in a post-processing step.

This parsing system obtains the best results for French, both for overall parsing and for MWE recognition, using a reparsing architecture that combines several parsers, with both pipeline architecture (MWE recognition followed by parsing), and joint architecture (MWE recognition performed by the parser).

6.13. Towards a French FrameNet

Participants: Marie-Hélène Candito, Marianne Djemaa, Benoît Sagot, Éric Villemonte de La Clergerie, Laurence Danlos.

The ASFALDA project [1] is a three-year project which started in October 2012, with the objective of building 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 [47], [2] which provides a structured set of prototypical situations, called *frames*, along with a semantic characterization of the participants of these situations (called *frame elements*, FEs). 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 the first year of the project, we focused on the first of these objectives. A team of 10 active members, from Alpage, the Laboratoire de Linguistique Formelle (LLF), the MELODI team (IRIT - Toulouse) and the CEA-List partners achieved :

- the delimitation and adaptation to French of a set of FrameNet frames, in order to cover a set of specific notional domains (commercial transaction, communication, cognitive positions, judgment/evaluation, temporal relations, spatial position, causality).
- and the semi-automatic construction of a French lexicon in which French lexical units are associated with frames

The current resource contains 110 frames, and roughly 2500 lexical units / frame pairs. The next phase consists in automatic pre-annotation of semantic annotations, that will serve as basis for the manual validation phase.

Note that a publication describing the project and these first achievements shall be presented at the LREC 2014 conference.

[1] https://sites.google.com/site/anrasfalda/
[2] https://framenet.icsi.berkeley.edu/

6.14. Modelisation of discourse structures with DSTAG

Participant: Laurence Danlos.

This work was done within the ANR Polymnie, in collaboration with Sylvain Pogodalla and Philippe de Groote from LORIA.

Neg-Raising (NR) verbs form a class of verbs with a clausal complement that show the following behavior: when a negation syntactically attaches to the matrix predicate, it can semantically attach to the embedded predicate. Such an implication does not always hold. Some contexts make it impossible to consider the negation as having scope over the embedded predicate only. This corresponds to the non-NR reading of the predicate.

We have developed and published [20] an account of NR predicates within Tree Adjoining Grammars (TAG) that relies on a Montague-like semantics for TAG. The different properties of NR predicates are rendered at different levels: the ambiguity of the readings is modeled by lexical ambiguity; the scoping and cyclicity properties are modeled through the lexical semantics and the higher-order interpretation of adjunction nodes; spurious am- biguities are avoided using fine-grained types for terms representing derivation trees. This provides us with a base layer where to account for interactions with discourse connectives and discourse representation represented in DSTAG.

6.15. Annotation of discourse structures on the FTB

Participants: Laurence Danlos, Margot Colinet.

With the aim of annotating the French TreeBank (FTB, already annotated for syntax) with discourse information, we have been working on the first step of the project, namely identify all the occurrences of discourse connectives in the FTB. This raises problems for lexemes which are ambiguous with a discourse usage and other uses. In collaboration with Mathilde Dargnat (ATILF) and Grégoire Winterstein, we have been working on the preposition *pour* (around 1500 occurrences) and the adverb *alors* (300 occurrences). This work is the basis for a future annotation manual.

In parallel, we have been working on adverbial discourse connectives and published on the topic [17]. This paper focuses on the following question: does the only syntactic argument of an adverbial discourse connective correspond to its second semantic argument? It shows that this is not always the case, which is a problem for the syntax-semantics interface. This interface brings us to distinguish two classes of adverbial connectives we sketch the study of.

6.16. Pairwise coreference models

Participant: Emmanuel Lassalle.

In collaboration with Pascal Denis (Magnet, Inria), we have proposed a new method for significantly improving the performance of pairwise coreference models [34]. Given a set of indicators, our method learns how to best separate types of mention pairs into equivalence classes for which we construct distinct classification models. In effect, our approach finds an optimal fea- ture space (derived from a base feature set and indicator set) for discriminating coreferential mention pairs. Although our approach explores a very large space of possible feature spaces, it remains tractable by exploiting the structure of the hierarchies built from the indicators.

In the framework of decision trees, this method can be seen as a pruning procedure and thus can be combined with different methods for expanding a decision tree. It can also be compared to polynomial kernels, but has the advantage of a lower computational complexity [21]. Our experiments on the CoNLL-2012 Shared Task English datasets (gold mentions) indicate that our method is robust relative to different clustering strategies and evaluation metrics, showing large and consistent improvements over a single pairwise model using the same base features. Our best system obtains a competitive 67.2 of average F1 over MUC, B3, and CEAF which, despite its simplicity, places it above the mean score of other systems on these datasets.

6.17. Identification of implicit discourse relations

Participant: Chloé Braud.

In collaboration with Pascal Denis (Magnet, Inria), we have developed a system for identifying "implicit" discourse relations (that is, relations that are not marked by a discourse connective) [33]. Given the little amount of available annotated data for this task, our system also resorts to additional automatically labeled data wherein unambiguous connectives have been suppressed and used as relation labels, a method introduced by Marcu and Echihabi (2002). As shown by Sporleder and Lascarides (2008) for English, this approach doesn't generalize well to implicit relations as annotated by humans. We have shown that the same conclusion applies to French due to important distribution differences between the two types of data. In consequence, we propose various simple methods, all inspired from work on domain adaptation, with the aim of better combining annotated data and artificial data. We have evaluated these methods through various experiments carried out on the ANNODIS corpus: our best system reaches a labeling accuracy of 45.6%, corresponding to a 5.9% significant gain over a system solely trained on manually labeled data.

ALPINES Team

6. New Results

6.1. Integral equations on multi-screens

We developed a new functional framework for the study of scalar wave scattering by objects, called multi-screens, that are arbitrary arrangements of thin panels of impenetrable materials. From a geometric point of view, multi-screens are a priori non-orientable non-Lipschitz surfaces. We use our new framework to study boundary integral formulations of the scattering by such objects.

6.2. Second-kind Galerkin boundary element method for scattering at composite objects

In the context of scattering of time-harmonic acoustic waves at objects composed of several homogeneous parts with different material properties, a novel second-kind boundary integral formulation for this scattering problem was proposed in [X. Claeys, A single trace integral formulation of the second kind for acoustic scattering, Report 2011-14, SAM, ETH Zürich]. We recasted it into a variational problem set in L2 and investigated its Galerkin boundary element discretization from a theoretical and algorithmic point of view. Empiric studies demonstrate the competitive accuracy and superior conditioning of the new approach compared to a widely used Galerkin boundary element approach based on a first-kind boundary integral formulation.

6.3. Instability phenomenon for a rounded corner in presence of a negative material

We studied a 2D transmission problem between a positive material and a negative material. In electromagnetics, this negative material can be a metal at optical frequencies or a negative metamaterial. We highlighted an unusual instability phenomenon in some configurations: when the interface between the two materials presents a rounded corner, it can happen that the solution depends critically on the value of the rounding parameter. To prove this result, we provided an asymptotic expansion of the solution, when it is well-defined, in the geometry with a rounded corner. Then, we demonstrated that the asymptotic expansion is not stable with respect to the rounding parameter. We also conducted obtained numerical experiments with finite element methods to validate these results.

6.4. Parallel design and performance of direction preserving preconditioners

In the context of preconditioned iterative methods, our work has focused on so called direction preserving preconditioners. In [9] we consider the parallel design and performance of nested filtering factorization (NFF), a multilevel parallel preconditioning technique for solving large sparse linear systems of equations by using iterative methods. NFF is based on a recursive decomposition that requires first to permute the input matrix, which can have an arbitrary sparsity structure, into a matrix with a nested block arrow structure. This recursive factorization is a key feature in allowing NFF to have limited memory requirements and also to be very well suited for hierarchical parallel machines. NFF is also able to preserve some directions of interest of the input matrix A. Given a set of vectors T which represent the directions to be preserved, the preconditioner M satisfies a right filtering property $MT = AT$. This is a property which has been exploited in different contexts, as multigrid methods [Brandt et al., 2011, SIAM J. Sci. Comput.], semiseparable matrices [Gu et al, 2010, SIAM J. Matrix Anal. Appl.], incomplete factorizations [Wagner, 1997, Numer. Math] , or nested factorization [Appleyard and Cheshire, 1983, SPE Symposium on Reservoir Simulation]. It is well known that for difficult problems with heterogeneities or multiscale physics, the iterative methods can converge very slowly, and this is often due to the presence of several low frequency modes. By preserving the directions

corresponding to these low frequency modes in the preconditioner, their effect on the convergence is alleviated and a much faster convergence is often observed. NFF can be seen as an extension of nested factorization that can be used for matrices with arbitrary sparsity structure and for which the computation can be performed in parallel. While the algebra of NFF has been introduced previously [Grigori et al, 2010, Inria tech. report], we relate the arithmetic complexity of NFF to the depth of recursion of its decomposition, and with our data distribution and implementation, we estimate its arithmetic and communication complexity. We also discuss the convergence of NFF on a set of matrices arising from the discretization of a boundary value problem with highly heterogeneous coefficients on three-dimensional grids. Our results show that on a $400 \times 400 \times 400$ regular grid, the number of iterations with NFF increases slightly while increasing the number of subdomains up to 2048. In terms of runtime performance on Curie, a Bullx system formed by nodes of two eight-core Intel Sandy Bridge processors, NFF scales well up to 2048 cores and it is 2.6 times faster than the domain decomposition preconditioner Restricted Additive Schwarz (RAS) as implemented in PETSc http://www.mcs.anl.gov/petsc/. The choice of the filtering vectors plays an important role in direction preserving preconditioners. There are problems for which we have prior knowledge of the near kernel of the input matrix, and this is indeed the case for the problems tested in this paper. They can also be approximated by using techniques similar to the ones used in deflation, however we do not discuss further this option here.

6.5. New resuls in communication avoiding algorithms for sparse linear algebra

In the context of sparse linear algebra algorithms, our recent results focus on two operations, incomplete LU factorization preconditioners and sparse matrix-matrix multiplication.

In [12] 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 was no communication avoiding preconditioner available 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 k-way or nested dissection. We show that the reordering does not affect the convergence rate of the ILU0 preconditioned system as compared to k-way 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.

In [6] we consider a fundamental problem in combinatorial and scientific computing, the sparse matrix-matrix multiplication problem. Obtaining scalable algorithms for this operations is difficult, since this operation has a poor surface to volume ratio, that is a poor data re-use. We consider that the input matrices are random, corresponding to Erdos-Renyi random graphs. We determine new lower bounds on communication for this case, in which we assume that the algorithm is sparsity independent, where the computation is statically partitioned to processors independent of the sparsity structure of the input matrices. We show in this paper that existing algorithms for sparse matrix-matrix multiplication are sub-optimal in their communication costs, and we obtain new algorithms which are communication optimal, communicating less than the previous algorithms and matching new lower bounds.

6.6. New resuls in communication avoiding algorithms for dense linear algebra

In the context of dense linear algebra algorithms, we have focused on two operations, LU factorization and rank revealing QR factorization.

In [4] we present block LU factorization with panel rank revealing pivoting (block LU_PRRP), a decomposition algorithm based on strong rank revealing QR panel factorization. Block LU_PRRP is more stable than Gaussian elimination with partial pivoting (GEPP), with a theoretical upper bound of the growth factor of $(1+\tau b)^{(n/b)-1}$, where b is the size of the panel used during the block factorization, τ is a parameter of the strong rank revealing QR factorization, n is the number of columns of the matrix, and for simplicity we assume that n is a multiple of b. We also assume throughout all the paper that $2 \leq b \leq n$. For example, if the size of the panel is $b = 64$, and $\tau = 2$, then $(1+2b)^{(n/b)-1} = (1.079)^{n-64} \ll 2^{n-1}$, where 2^{n-1} is the upper bound of the growth factor of GEPP. Our extensive numerical experiments show that the new factorization scheme is as numerically stable as GEPP in practice, but it is more resistant to pathological cases. The block LU_PRRP factorization does only $O(n^2 b)$ additional floating point operations compared to GEPP.

We also present block CALU_PRRP, a version of block LU_PRRP that minimizes communication, and is based on tournament pivoting, with the selection of the pivots at each step of the tournament being performed via strong rank revealing QR factorization. Block CALU_PRRP is more stable than CALU, the communication avoiding version of GEPP, with a theoretical upper bound of the growth factor of $(1+\tau b)^{\frac{n}{b}(H+1)-1}$, where H is the height of the reduction tree used during tournament pivoting. The upper bound of the growth factor of CALU is $2^{n(H+1)-1}$. Block CALU_PRRP is also more stable in practice and is resistant to pathological cases on which GEPP and CALU fail.

We have also introduced CARRQR (paper submitted to SIAM Journal on Matrix Analysis and Applications), a communication avoiding rank revealing QR factorization with tournament pivoting. 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 commmunication is a major bottleneck that highly impacts the performance of an algorithm.

6.7. Scalable Schwarz domain decomposition methods

Domain decomposition methods are, alongside multigrid methods, one of the dominant paradigms in contemporary large-scale partial differential equation simulation. A lightweight implementation [8] of a theoretically and numerically scalable preconditioner was developped in the context of overlapping methods. The performance of this work is assessed by numerical simulations executed on thousands of cores, for solving various highly heterogeneous elliptic problems in both 2D and 3D with billions of degrees of freedom. Such problems arise in computational science and engineering, in solid and fluid mechanics.
For example, in 3D, the initial highly heterogeneous problem of 74 million unknowns is solved in 200 seconds on 512 threads. Using 16384 threads, the problem is now made of approximately 2.3 billions unknowns, and it is solved in 215 seconds, which yields an efficiency of $\approx 90\%$. In 2D, the initial problem of 695 million unknowns is solved in 175 seconds on 512 threads. Using 16384 threads, the problem is now made of approximately 22.3 billions unknowns, and it is solved in 187 seconds, which yields an efficiency of $\approx 96\%$.

6.8. Schur domain decomposition methods

We have introduced spectral coarse spaces for the BDD and FETI methods in [5]. These coarse spaces are specifically designed for the two-level methods to be scalable and robust with respect to the coefficients in the equation and the choice of the decomposition. We achieve this by solving generalized eigenvalue problems on the interfaces between subdomains to identify the modes which slow down convergence. Theoretical bounds for the condition numbers of the preconditioned operators which depend only on a chosen threshold and the maximal number of neighbours of a subdomain were proved. For FETI there are two versions of the two-level method: one based on the full Dirichlet preconditioner and the other on the, cheaper, lumped preconditioner. Some numerical tests confirm these results.

6.9. Non conforming domain decomposition methods

We have designed and analyzed a new non-conforming domain decomposition method, named the NICEM method, based on Schwarz-type approaches that allows for the use of Robin interface conditions on non-conforming grids. The method is proven to be well posed. The error analysis is performed in 2D and in 3D for P1 elements. Numerical results in 2D illustrate the new method. This work is in collaboration with C. Japhet and Y. Maday.

6.10. Quadratic finite elements with non-matching grids for the unilateral boundary contact

We analyze in [3] a numerical model for the Signorini unilateral contact, based on the mortar blue method, in the quadratic finite element context. The mortar frame enables one to use non-matching grids and brings facilities in the mesh generation of different components of a complex system. The convergence rates we state here are similar to those already obtained for the Signorini problem when discretized on conforming meshes. The matching for the unilateral contact driven by mortars preserves then the proper accuracy of the quadratic finite elements. This approach has already been used and proved to be reliable for the unilateral contact problems even for large deformations. We provide however some numerical examples to support the theoretical predictions with FreeFem++ (http://www.freefem.org/ff++).

ANGE Team

6. New Results

6.1. Geophysical flows

6.1.1. A numerical scheme for the Saint-Venant equations

Participants: Emmanuel Audusse, Christophe Chalons [Univ. Versailles], Philippe Ung.

In order to improve the numerical simulations of the shallow-water equations, one has to face three important issues related to the well-balanced, positivity and entropy-preserving properties, as well as the ability to handle vacuum states. In that purpose, we propose a Godunov-type method based on the design of a three-wave Approximate Riemann Solver (ARS) which satisfies all aforementioned properties.

6.1.2. Two-phase flows

Participants: Frédéric Coquel [CNRS], Jean-Marc Hérard [EDF], Khaled Saleh [IRSN], Nicolas Seguin.

After having developed numerical schemes for models of compressible two-phase flows [17], [19], we have proven some fundamental properties of these systems: symmetrizability and (non strict) convexity of the entropy [18]. This enables us now to address the well-posedness of these models when the relaxation terms are included.

6.1.3. Non-hydrostatic models

Participants: Marie-Odile Bristeau, Dena Kazerani, Anne Mangeney, Jacques Sainte-Marie, Nicolas Seguin.

The objective is to derive a model corresponding to a depth averaged version of the incompressible Euler equations with free surface. We have already contributed to this subject but the obtained results extend previous ones [29] in several directions:

- the derivation of the model is more rigorous and follows the entropy-based moment closures proposed in [28],
- the properties of the model and especially its connections with Green-Nagdhi model have been investigated,
- a family of analytical solutions for the proposed model have been obtained.

These analytical solutions emphasize the non-hydrostatic effects appearing for large slope variations.

6.1.4. Fluids with complex rheology

Participants: Anne Mangeney, Jacques Sainte-Marie.

We have been able

- to develop detection, characterization and localisation methods applicable to the seismic signals generated by rockfalls and thus to analyse ths spatio-temporal change of rockfall localisation and properties during several years, making it possible to show how rockfalls can be used as a precursor of volcanic activity,
- to propose an empirical “universal” law describing friction weakening in landslides over a broad range of volumes and geological contexts,
- to propose a new debris flow model with an energy balance,
- show the existence of a slow propagation phase in granular flows, playing a key role in their dynamics and in erosion processes.

6.1.5. Dynamics of sedimentary river beds with stochastic fluctuations

Participants: Emmanuel Audusse, Philippe Ung.

The Exner equation is a coarse model for the dynamics of sedimentary river beds, derived using both many heuristics and empirism. Though, it is also quite practical for hydraulic engineering applications, and efficient enough in numerous situations. Our goal in this work is to improve the model by including some effects that have been neglected so far in the heuristics. In particular, inline with other current research directions in the field, we study the possibility of introducing some stochasticity in the model. To this end, we suggest to numerically experiment some recently proposed variations of the Exner equation based on the introduction of stochastic fluctuations within the standard formulation.

This project has been the subject of a study during the 2013 session of the CEMRACS.

6.2. Ecology and sustainable energies

6.2.1. Hydrodynamic-biology coupling

Participants: Olivier Bernard [Inria BIOCORE], Anne-Céline Boulanger, Marie-Odile Bristeau, Raouf Hamouda, Jacques Sainte-Marie.

An important part of our research activity is built around a biological and industrial problem: the simulation of the coupling of hydrodynamics and biology in the context of industrial microalgae culture in outdoor raceways. The numerical modelling is adressed with the use of a multilayer vertical discretization of hydrostatic Navier-Stokes equations coupled with a light sensitive Droop model. Numerically, kinetic schemes allow for the development of efficient, positivity preserving, well balanced and entropy satisfying schemes. Simulations are carried out in 2D and 3D [1]. From a practical point of view, this model is capable of accounting for the utility of a paddlewheel and exhibits Lagrangian trajectories underwent by algae. Hence providing hints on the light history of algae in the pond, which is a key information to biologists, since it enables them to adapt their phytoplankton growth models to those particular, non natural conditions.

6.3. Coupling methods

6.3.1. Data assimilation for conservation laws associated with kinetic description

Participants: Anne-Céline Boulanger, Philippe Moireau [Inria M3DISIM], Jacques Sainte-Marie.

In order to take advantage of the kinetic description of conservation laws already used for the building of efficient schemes, an innovative data assimilation method for hyperbolic balance laws based in a Luenberger observer on the kinetic equation is developed. It provides a nice theoretical framework for scalar conservation laws, for which we study the cases of complete observations, partial observations in space, in time, and noisy observations. As far as systems are concerned, we focus on the Saint-Venant system, which is hyperbolic, nonlinear and has a topographic source term. We build an observer based only on water depths measurements. Numerical simulations are provided in the case of scalar laws and systems, in one and two dimensions, which validate the efficiency of the method [14].

6.3.2. Mach-parametrized flows

Participants: Stéphane Dellacherie [CEA], Bruno Després [UPMC Paris 6], Yohan Penel.

In order to enrich the modelling of fluid flows, we investigate in this paper a coupling between two models dedicated to distinct regimes. More precisely, we focus on the influence of the Mach number as the low Mach case is known to induce theoretical and numerical issues in a compressible framework. A moving interface is introduced to separate a compressible model (Euler with source term) and its low Mach counterpart through relevant transmission conditions. A global steady state for the coupled problem is exhibited. Numerical simulations are then performed to highlight the influence of the coupling by means of a robust numerical strategy [20].

6.3.3. Error analysis in a coupling strategy

Participants: Clément Cancès [UPMC Paris 6], Frédéric Coquel [CNRS], Edwige Godlewski, Hélène Mathis [Univ. Nantes], Nicolas Seguin.

We have proposed in a simplified framework an error analysis for an adaptive method which automatically selects the optimal model to use, the choice being between a reference model and an associated simplified one, see [15]. In particular, we are able to balance the thickness of the coupling buffer zone with the threshold on the modelling error which appears when introducing the coarse model.

6.4. Software development and assessments

6.4.1. Analytical solutions for the incompressible Euler system

Participants: Anne-Céline Boulanger, Marie-Odile Bristeau, Jacques Sainte-Marie.

We have proposed in [5] a large set of analytical solutions (FRESH-ASSESS) for the hydrostatic incompressible Euler system in 2d and 3d. These solutions mainly concern free surface flows but partially free surface flows are also considered. These analytical solutions can be especially useful for the validation of numerical schemes.

6.4.2. Software

Several tasks have been achieved in the FRESHKISS3D software (§ 5.1):

- First tests with a uniform pression before moving to the variable case;
- Rethinking of the C++ code with an object-oriented rewriting which provides a better memory management;
- Automatic boundary conditions handling in the case of a fluid/solid transition;
- New computations of the particule trajectories when they leave out the domain by means of directional interpolation procedures;
- Achievement of the 2nd-order space accuracy;
- Taking into account the wind.

AOSTE Project-Team

6. New Results

6.1. Process Networks with routing for parallel architectures

Participants: Robert de Simone, Emilien Kofman, Jean-Vivien Millo.

In the past we developed a dedicated Process Network (PN) formalism with explicit static switching/routing schemes for data flow. This year we considered the practical use of our formalism to model data-streams in specific applicative contexts.
In a first direction we considered the case of stencil algorithms, usually modeled with cellular automata (CA) (as in heat or gas propagation models for instance). In that case, the application itself is modeled in a way strongly similar to a physical architecture consisting of a regular mesh/array of parallel processors (MPPA). Mapping can seem to be straighforward then, *safe that* the neighborhood and connection topology may differ from the CA model to the MPPA. Our results consider efficient routing and propagation schemes on a given MPPA interconnect fabric, so as to match all-to-all broadcast paterns up to a given distance (on the CA topology). They are described in [20], and were implemented on Kalray MPPA256 prototype architecture.
A similar modeling effort was conducted, this time on FFT algorithm models (again described as parallel pipe-lined tasks). Again switching/routing schemes were provided in our formal PN model to map the virtual logical dependences onto concrete connection patterns in a MPPA256 model. This was the subject of Emilien Kofman internship, of which preliminary results were presented in a junior workshop [36].

6.2. Formal analysis of MARTE Time Model and CCSL

Participants: Frédéric Mallet, Robert de Simone, Yuliia Romenska, Jean-Vivien Millo, Ling Yin.

We have worked on building analysis methods and tools for running exhaustive analyses on MARTE/CCSL specifications. This was done by endowing CCSL with a State-Based semantics [51]. Each operator is described as a boolean state machine, some operators require an infinite number of states. When this is the case we rely on a lazy representation technique to capture symbolically the infinite number of states [45]. The semantics of a CCSL specification is then expressed as the synchronized product of the (infinite) state machines for each operator. Even though the operators are infinite, their composition can sometimes be bounded. When the synchronized product has only a finite number of reachable states, it is said to be safe. We have identified a set of representative and frequently used examples where this is the case [38]. When the product is not finite, our (semi-)algorithm to build the product does not terminate, therefore it is important to be able to know in advance whether or not the product is safe. We have thus proposed an algorithm to decide whether a CCSL specification is safe [37]. It relies on an intermediate representation called Clock Causality Graph and uses results from marked graph theory.

Building the product for a CCSL specification is exponential in the number of clocks and is not practical for large specifications. So, to avoid building explicitly the product we have proposed another technique to explore symbolically the state-space of a CCSL specification [49]. This relies on a liveness condition where no conflict may prevent an infinite clock from ticking infinitely often. Branches that may lead to states where an infinite clock dies are pruned by a fix-point algorithm.

These two solutions focus on the logical and discrete aspects of MARTE/CCSL, which was devised to unify logical and physical time constraints. An attempt to support verification of the physical time constraints of MARTE/CCSL was conducted through the use of UppAal timed automata and model-checker [46]. The proposed technique combines the logical clocks of CCSL with the real-valued clocks of timed automata. Synchronous/Polychronous aspects are solved with TimeSquare 5.1 while the UppAal model-checker is used to explore the space derived from the real-valued clocks.

6.3. Logical time in Model-Driven Engineering of embedded systems

Participants: Frédéric Mallet, Julien Deantoni, Robert de Simone, Marie-Agnès Peraldi Frati, Matias Vara-Larsen, Arda Goknil.

In the context of our approach based on logical time to specify causalities and synchronizations on models, 3.2 , we developed an extension of the OMG OCL Object Constraint Language. Named ECL (Event Constraint Language) it provides such specifications of causalitity and synchronization at syntactic language level, which enabled then automatic generation of semantic logical time constraints for any model that conforms the language.
This year, we extended to a new challenge, using logical time constraints to coordinate models of *several distinct* languages used jointly for a large heterogeneous system description. This work is reported in [25], [52].
It was illustrated in practice in the automotive domain by coordinating together the Timed Augmented Description Language (TADL2) and the EAST-ADL language [34], [32] (the formalisms are rather similar, but still with clear distinctions at places) .
Finally, we proposed a pattern to assemble the (possibly concurrent) semantics of a language associating our logical time constraints (based on pure clocks) with a syntactic action language (providing behavior content). By reifying events and constraints, this specification of the semantics is amenable to its composition [25]. Such approach has been, again, recently used for a first attempt to coordinate distinct behavioral models [47].

As part of our collaboration in the DAESD associated-team with ECNU Shone-SEI in Shanghai we studied the coupling of discrete-logical with continuous-physical time models, ending with a proposal of Hybrid MARTE statecharts [19] specified in a style much like a combinaison of MARTE state diagrams and timed automata.

In another setting we presented a new model of scenarios [21], dedicated to the specification and verification of system behaviours in the context of software product lines (SPL). The formalism uses the logical time modeling aproach, with a strong link to synchronous semantics. We draw our inspiration from some techniques that are mostly used in the hardware community, and we show how they could be applied to the verification of software components and product line variability. We point out the benefits of synchronous languages and models to bridge the gap between both worlds.

6.4. Multiview modeling and power intent in Systems-on-chip

Participants: Carlos Gomez Cardenas, Ameni Khecharem, Emilien Kofman, Frédéric Mallet, Julien Deantoni, Robert de Simone.

Power models for embedded architectures (where power consumption is highly constrained) provide an ideal example of a non-functional modeling framework with strong interactions with the functional and performance models: more speed in computation comes at the cost of larger energy consumption. There was also a demand for a framework allowing combinaison of models, each representing a distinct view of the system. We demonstrated as part of the HeLP ANR project 8.2.1.1 , followed by the newly started HOPE ANR project 8.2.1.2 , how such multiview modeling could be done, and how it could be connected down to more concrete simulation code or model, as in SystemC, Docea Power AcePlorer, or Scilab code. The multiview modeling applied to power intent and power managers was described in [35], and led to the PhD defense of Carlos Gomez Cardenas in December 2013 [16].

6.5. Performance variability analysis on manycore architectures

Participants: Sid Touati, Amin Oueslati, Franco Pestarini, Robert de Simone, Emilien Kofman.

In the context of the collaboration with Kalray (see 7.1.1), we conducted a systematic benchmarking campaign to test the stability (or low variability) of the performances of the MPPA256 prototype manycore processor. We first addressed issues of memory access and network latency, then programmed a distributed verson of the classical ALL_PAIRS_SHORTEST_PATH parallel algorithm with an hybrid OpenMP/MPI style. This was the objectif of Amin Oueslati Master2 internship. Results were encouraging, and showed stability of performance over a large set of runs.

This work is currently extended during the International Internship grant of Franco Pescarini. Specific on-chip communication modes offered by the MPPA256 processor (namely *portal* and *channel* communication modes) are being extensively benchmarked. Results show time predictability on the case of light on-chip communication traffic, but stability gets degraded as performance decreases in presence of heavy traffic and congestion (various runs show quite different execution time).
In another effort we conducted during the internship period of Emilien Kofman an experiment on MPPA256 quite similar to the work conducted as part of the collaboration with Kontron (see 7.1.3), exploring various mapping options of FFT algorithm variants, with the goal of figuring how to best map (in the future) several such algorithms onto the computation fabric of the many-cores available.

6.6. Off-line (static) mapping of real-time applications onto NoC-based many-cores

Participants: Thomas Carle, Manel Djemal, Dumitru Potop Butucaru, Robert de Simone, Zhen Zhang.

Modern computer architectures are increasingly relying on multi-processor systems-on-chips (MPSoCs, also called chip-multiprocessors), with data transfers between cores and RAM banks managed by on-chip networks (NoCs). This reflects in part a convergence between embedded, general-purpose PC, and high-performance computing (HPC) architecture designs. In past years we have identified and compared the hardware mechanisms supporting precise timing analysis and efficient resource allocation in existing NoCs. We determined that the NoC should ideally provide the means of enforcing a global communications schedule that is computed off-line and which is synchronized with the scheduling of computations on CPU cores (and we have built such a NoC).

This year we have focused on the problem of mapping applications onto NoC-based MPSoCs (discussed in this section) and on the associated problem of timing analysis of the resulting parallel implementations (discussed in section 6.7). On-chip networks used in MPSoCs pose significant challenges to both on-line and off-line real-time scheduling approaches. They have large numbers of potential contention points, have limited internal buffering capabilities, and network control operates at the scale of small data packets. Therefore, precise schedulability analysis requires scalable algorithms working on hardware models with a level of detail that is unprecedented in real-time scheduling.

We considered an off-line scheduling approach, and we targeted massively parallel processor arrays (MPPAs), which are MPSoCs with large numbers (hundreds) of processing cores. We proposed a novel allocation and scheduling method capable of synthesizing such global computation and communication schedules covering all the execution, communication, and memory resources in an MPPA. To allow an efficient use of the hardware resources, our method takes into account the specificities of MPPA hardware and implements advanced scheduling techniques such as pre-computed preemption of data transmissions and pipelined scheduling.

Our method has been implemented within the Lopht tool presented in section 5.4 , and first results are presented in [54]. One of the objectives of the collaboration with Kalray SA is the evaluation of the possibility of porting Lopht onto the Kalray MPPA platform.

6.7. WCET estimation for parallel code

Participant: Dumitru Potop Butucaru.

This is joint work with Isabelle Puaut, Inria, EPI ALF.

Classical timing analysis techniques for parallel code isolate micro-architecture analysis from the analysis of synchronizations between cores by performing them in two separate analysis phases (WCET – worst-case execution time – and WCRT – worst-case response time analyses). This isolation has its advantages, such as a reduction of the complexity of each analysis phase, and a separation of concerns that facilitates the development of analysis tools. But isolation also has a major drawback: a loss in precision which can be significant. To consider only one aspect, to be safe the WCET analysis of each synchronization-free sequential code region has to consider an undetermined micro-architecture state. This may result in overestimated WCETs, and consequently on pessimistic execution time bounds for the whole parallel application.

The contribution of this work [56], [44] is an *integrated* WCET analysis approach that considers at the same time micro-architectural information and the synchronizations between cores. This is achieved by extending a state-of-the-art WCET estimation technique and tool to manage synchronizations and communications between the sequential threads running on the different cores. The benefits of the proposed method are twofold. On the one hand, the micro-architectural state is not lost between synchronization-free code regions running on the same core, which results in tighter execution time estimates. On the other hand, only one tool is required for the temporal validation of the parallel application, which reduces the complexity of the timing validation toolchain.

Such a holistic approach is made possible by the use of deterministic and composable software and hardware architectures (many-cores with no cache sharing and time-predictable interconnect, static assignment of the code and data to the memory banks). Such code can be written by hand or automatically synthesized using the Lopht tool 5.4 or other automatic parallelization techniques.

6.8. Real-time scheduling and code generation for time-triggered platforms

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

We have continued this year the work on real-time scheduling and code generation for time-triggered platforms. This work was mainly carried out as part of a bilateral collaboration with Astrium Space Transportation (now part of Airbus Defence and Space), which co-funded with the CNES the post-doctorate of Raul Gorcitz (started in September).

The work focused this year on the improvement of the real-time scheduling and code generation (the PhD work of T. Carle), and on determining their adequacy to Astrium's industrial needs (the post-doc of Raul Gorcitz). We have improved our specification, mapping, and code generation technique at all levels. We have extended the Lopht tool to allow automatic mapping and code generation for single-processor and multi-processor partitioned targets (using an ARINC 653-compliant OS).

6.9. Uniprocessor Real-Time Scheduling

Participants: Yves Sorel, Falou Ndoye, Daniel de Rauglaudre.

6.9.1. Formal Proofs of Uniprocessor Real-Time Scheduling Theorems

We continued writing a monograph about three formal proofs, done in 2011/2012, in Coq on scheduling of fixed priority real-time preemptive tasks: one about the scheduling conditions of strict periodicity and two about the worst response time in the case of preemptive deadline monotonic scheduling. This document contains about 120 pages for the moment.

6.9.2. Real-Time Scheduling with Exact Preemption Cost

We proposed a new schedulability condition for dependent tasks executed on a uniprocessor which takes into account the exact preemption cost. Unlike the work presented in [10] which achieves that goal only for fixed priority tasks, our schedulability condition considers fixed as well as dynamic priorities tasks. Thus, we can overcome priority inversions involved by data dependent tasks. The schedulability analysis based on this schedulability condition led to an off-line scheduler [42] described by a scheduling table. Therefore, we have proposed an on-line time-trigger scheduler which implements this scheduling table. Compared to classical on-line schedulers, the proposed approach has two benefits. On the one hand the cost of the task selection amounts only to read the task to be executed in the scheduling table built off-line, rather than using on-line a scheduling algorithm like RM, DM, EDF, etc. On the other hand this cost is fixed since it does not depend on the number of ready tasks. In addition, with our on-line scheduler we do not need to synchronize, on-line, the utilization of the shared memory data, due to dependences, because this synchronization is performed during the off-line schedulability analysis.

6.10. Multiprocessor Real-Time Scheduling

Participants: Yves Sorel, Laurent George, Dumitru Potop-Butucaru, Falou Ndoye, Aderraouf Benyahia, Cécile Stentzel, Meriem Zidouni.

6.10.1. Multiprocessor Partitioned Scheduling with Exact Preemption Cost

We finalized the work started in previous years on multiprocessor scheduling of preemptive independent real-time tasks with exact preemption cost [43].

This year we proposed a heuristic for the multiprocessor scheduling of preemptive dependent real-time tasks with exact preemption cost. We chose the partitioned approach that avoids migration of tasks and allows the utilization of the uniprocessor schedulability condition, previously proposed, that takes into account the exact preemption cost. In addition, this schedulability condition takes into account the inter-processor communications and guarantees that no data is lost. The result of such an off-line scheduling provided by the heuristic, is a scheduling table for every processor which includes also inter-processor communication tasks. We compared our multiprocessor scheduling heuristic with a Branch & Bound exact algorithm using the same schedulability condition. Our heuristic provides similar results and is very much faster.

6.10.2. Multiprocessor Semi-Partitioned Mixed Criticality Scheduling

We mainly focused on the mixed criticality scheduling problem applied to semi-partitioned scheduling considering a static pattern of migration for jobs. We have studied this problem in the context of Mixed Criticality (MC) scheduling, a promising approach that can be used to take into account applications of different criticality levels on the same platform. The goal of MC approach is to better utilize computing resources by allowing low criticality tasks to execute in conjunction with high criticality tasks when the system criticality is not high.

6.10.3. Gateway with Modeling Languages for Certified Code Generation

This work was carried out in the P FUI project 8.2.2 . We defined a SynDEx UML profile for functional specifications. We developed a gateway between the P pivot formalism and SynDEx. This gateway deals with the data-flow modeling part of the P formalism which is compliant with the Simulink subset blocks supported by the P project, except for the IF, FOR, MERGE and MUX blocks. Presently, we enhance the gatetway to include these blocks and we colloborate with the other partners to define the architectural part of the P formalism. This part is intended to replace the non functional specifications, presently described with the UML profile MARTE (Modeling and Analysis of Real-Time Embedded Systems).

6.10.4. SynDEx updates with new results

We released an alpha version of SynDEx V8. This version is based on a new textual language whose compiler may be launched with commandes-lines featuring various options. In Syndex V8, the adequation heuristic which performs the multiprocessor real-time schedulability analysis on multi-periodic applications, is based on the theorems and algorithms provided in the Mohamed Marouf's thesis defended last year in the team. These algorithms have been deeply improved for better consideration of data dependencies in the case of multiprocessor architectures. On the other hand, the new heuristic generates a scheduling table composed of, in addition to the usual permanent phase, a transient phase that takes into account the distribution constraints defined by the user in the multi-periodic applications as well as in the mono-periodic applications.

6.11. Probabilistic Real-Time Systems

Participants: Liliana Cucu-Grosjean, Adriana Gogonel, Codé Lo, Dorin Maxim, Cristian Maxim.

The adventof complex hardware, in response to the increasing demand for computing power in next generation systems, exacerbates some of the limitations of static timing analysis for the estimation of the worst-case execution time (WCET) estimation. In particular, the effort of acquiring (1) detail information on the hardware to develop an accurate model of its execution latency as well as (2) knowledge of the timing behaviour of the program in the presence of varying hardware conditions, such as those dependent on the history of previously executed instructions. These problems are also known as the timing analysis walls. The probabilistic timing analysis, a novel approach to the analysis of the timing behaviour of next-generation real-time embedded systems, provides answers to timing analysis walls. In [17], [48], [31] timing analysis attacks the timing analysis walls. We have also presented experimental evidence that shows how probabilistic timing analysis reduces the extent of knowledge about the execution platform required to produce probabilistically-safe and tight WCET estimations.

Based on existing estimations of WCET or minimal inter-arrival time, one may propose different probabilistic schedulability analyses [39]. These results were reported in the (PhD thesis of Dorin Maxim, mostly conducted in the Inria TRIO team (before its completion and the move to Aoste in Sept 2013).

2013 was also the year when through several invited talks [26], [28], [27], we had the opportunity to underline historical misunderstandings on probabilistic real-time systems. The most common is related to the notion of independence that is used with a wrong meaning by different papers.

ARAMIS Team

6. New Results

6.1. Spatial and anatomical regularization of SVM

Participants: Rémi Cuingnet, Joan Glaunès, Marie Chupin, Habib Benali, Olivier Colliot [Correspondant].

We developed a general framework to introduce spatial and anatomical priors in SVM for brain image analysis based on regularization operators. A notion of proximity based on prior anatomical knowledge between the image points is defined by a graph (e.g. brain connectivity graph) or a metric (e.g. Fisher metric on statistical manifolds). A regularization operator is then defined from the graph Laplacian, in the discrete case, or from the Laplace-Beltrami operator, in the continuous case. The regularization operator is then introduced into the SVM, which exponentially penalizes high frequency components with respect to the graph or to the metric and thus constrains the classification function to be smooth with respect to the prior. It yields a new SVM optimization problem whose kernel is a heat kernel on graphs or on manifolds. We then present different types of priors and provide efficient computations of the Gram matrix. The proposed framework is finally applied to the classification of brain magnetic resonance (MR) images (based on gray matter concentration maps and cortical thickness measures) from 137 patients with Alzheimer's disease and 162 elderly controls. The results demonstrate that the proposed classifier generates less-noisy and consequently more interpretable feature maps (Figure 1) with high classification performances.

More details in [4].

6.2. Segmentation of the hippocampus in neurodegenerative dementias

Participants: Leonardo Cruz de Souza, Marie Chupin, Maxime Bertoux, Stéphane Lehéricy, Bruno Dubois, Foudil Lamari, Isabelle Le Ber, Michel Bottlaender, Olivier Colliot [Correspondant], Marie Sarazin.

Our team develops various applications of our automatic segmentation method SACHA to neurological disorders, in particular in neurodegenerative dementias. This research is done in close collaboration with IM2A (Institut de la Mémoire et de la Maladie d'Alzheimer, Bruno Dubois and Marie Sarazin) at Pitié-Salpêtrière hospital.

We previously showed that automatic hippocampal segmentation can discriminate patients with Alzheimer's disease (AD) from elderly control subjects, with high sensitivity and specificity. In patients with Alzheimer's disease, we further studied the relationship between hippocampal atrophy and memory deficits. We also showed that hippocampal volume loss is correlated to tau and hyperphosphorylated tau levels measured in the cerebro-spinal fluid (CSF) but not with $A\beta_4 2$ levels.

Here, our objective was to study the ability of hippocampal volumetry (HV) to differentiate between two neurodegenerative dementias: Alzheimer's disease (AD) and fronto-temporal dementia (FTD). Seventy-two participants were included: 31 AD patients with predominant and progressive episodic memory deficits associated with typical AD cerebrospinal fluid (CSF) profile and/or positive amyloid imaging (PET with 11C-labeled Pittsburgh Compound B [PiB]), 26 patients with behavioral variant FTD (bvFTD) diagnosed according to consensual clinical criteria and with no AD CSF profile, and 15 healthy controls without amyloid retention on PiB-PET exam. HV were segmented with our automated method and were normalized to total intracranial volume (nHV). Significant reductions in HV were found in both AD and bvFTD patients compared with controls, but there were no significant difference between AD and bvFTD patients. Mean nHV distinguished normal controls from either AD or bvFTD with high sensitivity (80.6% and 76.9%, respectively) and specificity (93.3% for both), but it was inefficient in differentiating AD from bvFTD (9.7% specificity). There was no difference in the clinical and neuropsychological profiles according to HV in bvFTD and AD patients. In conclusion, when considered alone, measures of HV are not good markers to differentiate AD from bvFTD. Hippocampal sclerosis associated with FTD may explain the high degree of overlap in nHV between both groups.

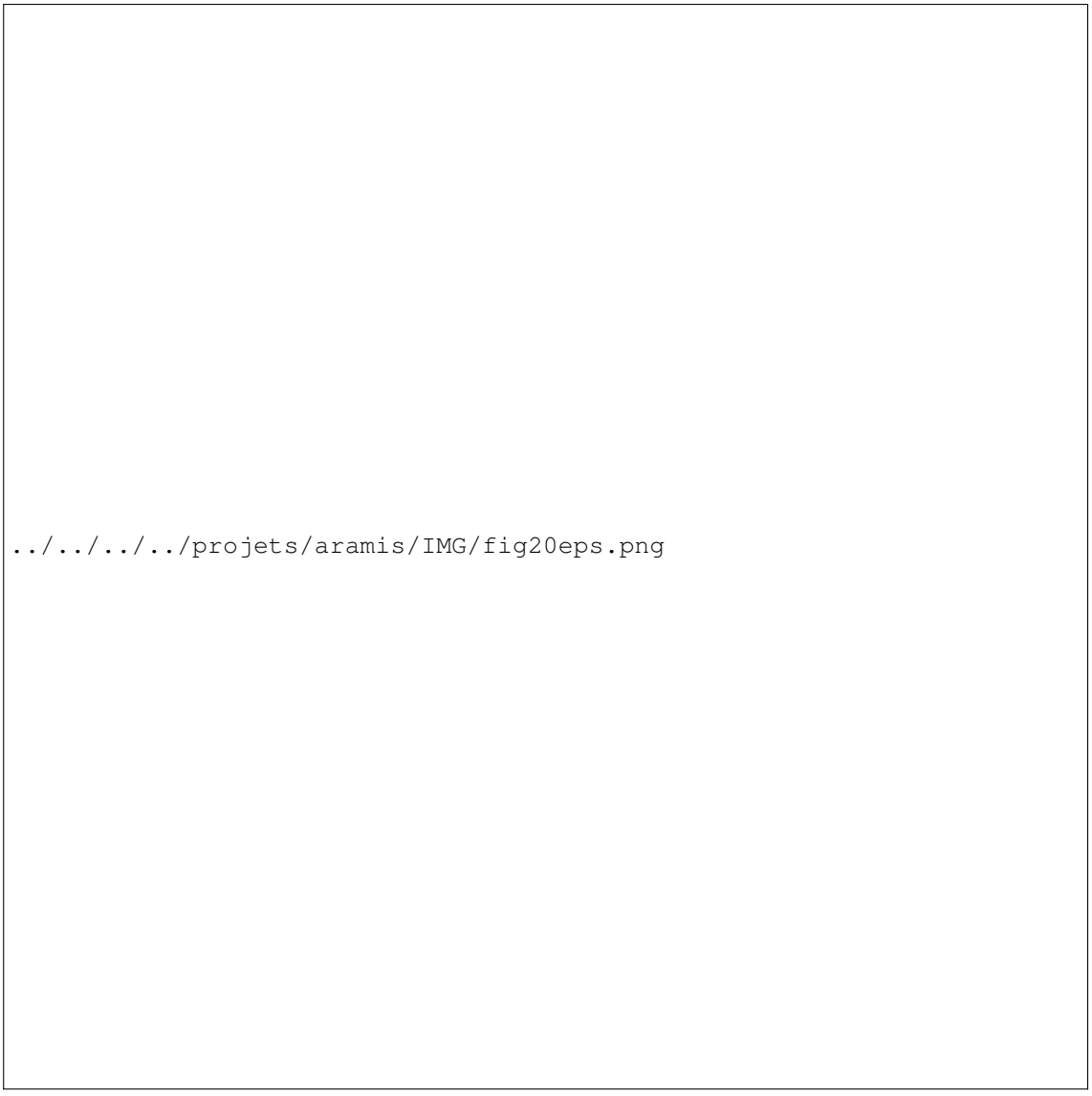


Figure 1. Anatomical regularization of support vector machines for automatic classification of patients with Alzheimer's disease. The figure displays the normalized vector orthogonal to the optimal margin hyperplane, for increasing levels of regularization.

More details in [5].

6.3. Diffeomorphic Iterative Centroids for Template Estimation on Large Datasets

Participants: Claire Cury [Correspondant], Joan Glaunès, Olivier Colliot.

A common approach for analysis of anatomical variability relies on the estimation of a template representative of the population. The Large Deformation Diffeomorphic Metric Mapping is an attractive framework for that purpose. However, template estimation using LDDMM is computationally expensive, which is a limitation for the study of large datasets. We proposed an iterative method which quickly provides a centroid of the population in the shape space. This centroid can be used as a rough template estimate or as initialization of a template estimation method. The approach was evaluated on datasets of real and synthetic hippocampi segmented from brain MRI. The results showed that the centroid is correctly centered within the population and is stable for different orderings of subjects. When used as an initialization, the approach allows to substantially reduce the computation time of template estimation.

More details in [30].

6.4. Sparse Adaptive Parameterization of Variability in Image Ensembles

Participants: Stanley Durrleman [Correspondant], Sarang Joshi, Stéphanie Allassonnière.

We introduce a new parameterization of diffeomorphic deformations for the characterization of the variability in image ensembles. Dense diffeomorphic deformations are built by interpolating the motion of a finite set of control points that forms a Hamiltonian flow of self-interacting particles. The proposed approach estimates a template image representative of a given image set, an optimal set of control points that focuses on the most variable parts of the image, and template-to-image registrations that quantify the variability within the image set. The method automatically selects the most relevant control points for the characterization of the image variability and estimates their optimal positions in the template domain. The optimization in position is done during the estimation of the deformations without adding any computational cost at each step of the gradient descent. The selection of the control points is done by adding a L^1 prior to the objective function, which is optimized using the FISTA algorithm.

Related publication: [12]

6.5. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data

Participants: Stanley Durrleman [Correspondant], Xavier Pennec, Alain Trouvé, José Braga, Guido Gerig, Nicholas Ayache.

We introduce a comprehensive framework for the statistical analysis of longitudinal shape data. The proposed method allows the characterization of typical growth patterns and subject-specific shape changes in repeated time-series observations of several subjects. This can be seen as the extension of usual longitudinal statistics of scalar measurements to high-dimensional shape or image data.

The method is based on the estimation of continuous subject-specific growth trajectories and the comparison of such temporal shape changes across subjects. Differences between growth trajectories are decomposed into morphological deformations, which account for shape changes independent of time, and time warps, which account for different rates of shape changes over time.

Given a longitudinal shape data set, we estimate a mean growth scenario representative of the population, and the variations of this scenario both in terms of shape changes and in terms of change in growth speed. Then, intrinsic statistics are derived in the space of spatiotemporal deformations, which characterize the typical variations in shape and in growth speed within the studied population. They can be used to detect systematic developmental delays across subjects.

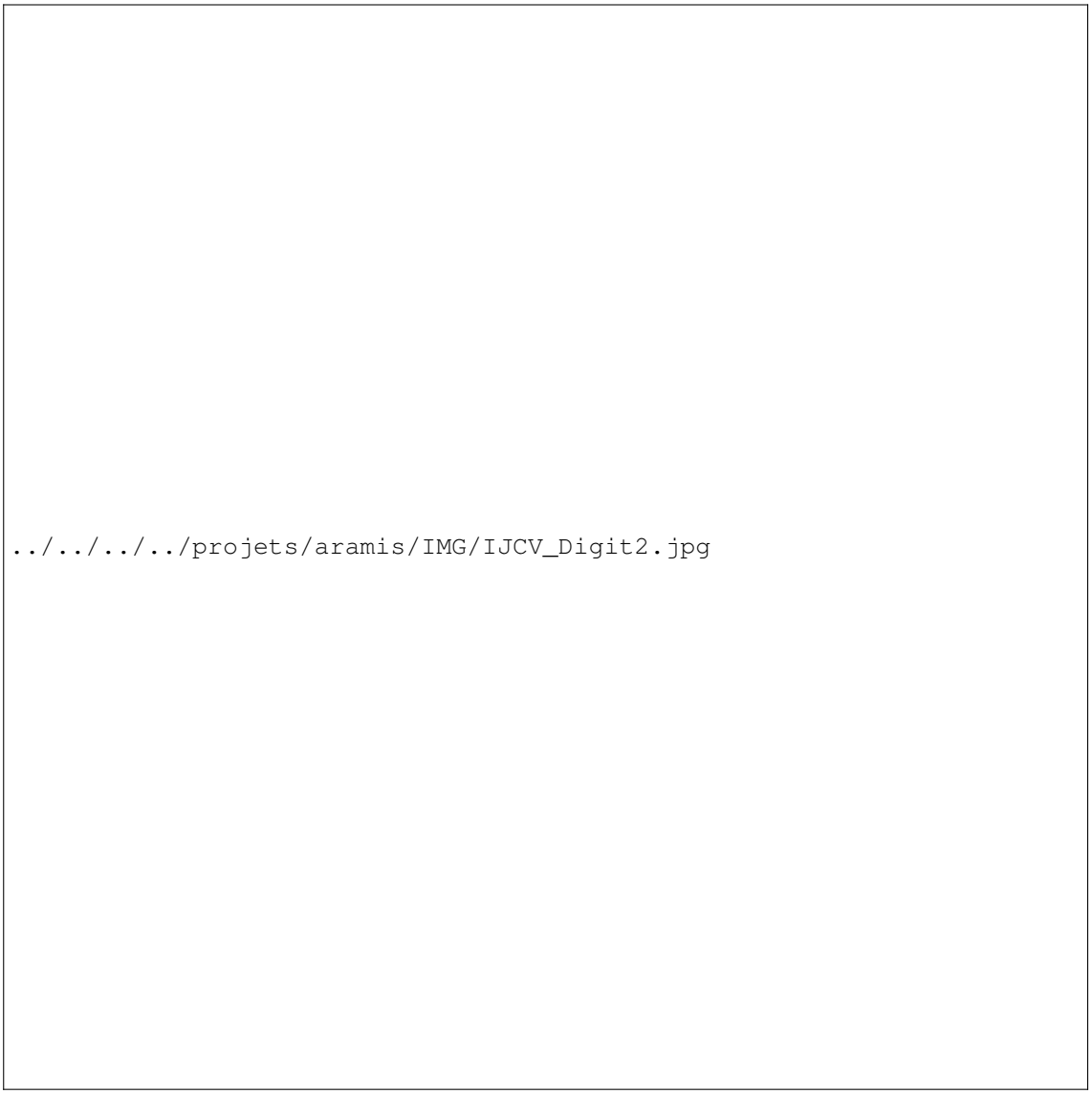


Figure 2. Left: template image estimated from 20 images of the US postal database. Momentum vectors are placed at the most variable places and paramterize mappings from the template to each image in the data set. Right: sample images from the data set (top) and template image deformed to match the corresponding sample image (bottom)

In the context of neuroscience, we apply this method to analyze the differences in the growth of the hippocampus in children diagnosed with autism, developmental delays and in controls. Result suggest that group differences may be better characterized by a different speed of maturation rather than shape differences at a given age. In the context of anthropology, we assess the differences in the typical growth of the endocranium between chimpanzees and bonobos. We take advantage of this study to show the robustness of the method with respect to change of parameters and perturbation of the age estimates.

Related publication: [13]

6.6. Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes

Participants: Pietro Gori [Correspondant], Olivier Colliot, Yulia Worbe, Linda Marrakchi-Kacem, Sophie Lecomte, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

We propose a Bayesian framework for multi-object atlas estimation based on the metric of currents which permits to deal with both curves and surfaces without relying on point correspondence. This approach aims to study brain morphometry as a whole and not as a set of different components, focusing mainly on the shape and relative position of different anatomical structures which is fundamental in neuro-anatomical studies. We propose a generic algorithm to estimate templates of sets of curves (fiber bundles) and closed surfaces (sub-cortical structures) which have the same "form" (topology) of the shapes present in the population. This atlas construction method is based on a Bayesian framework which brings to two main improvements with respect to previous shape based methods. First, it allows to estimate from the data set a parameter specific to each object which was previously fixed by the user: the trade-off between data-term and regularity of deformations. In a multi-object analysis these parameters balance the contributions of the different objects and the need for an automatic estimation is even more crucial. Second, the covariance matrix of the deformation parameters is estimated during the atlas construction in a way which is less sensitive to the outliers of the population.

Related publication: [33]

6.7. Geodesic regression of shape and image data

Participants: James Fishbaugh [Correspondant], Marcel Prastawa, Guido Gerig, Stanley Durrleman.

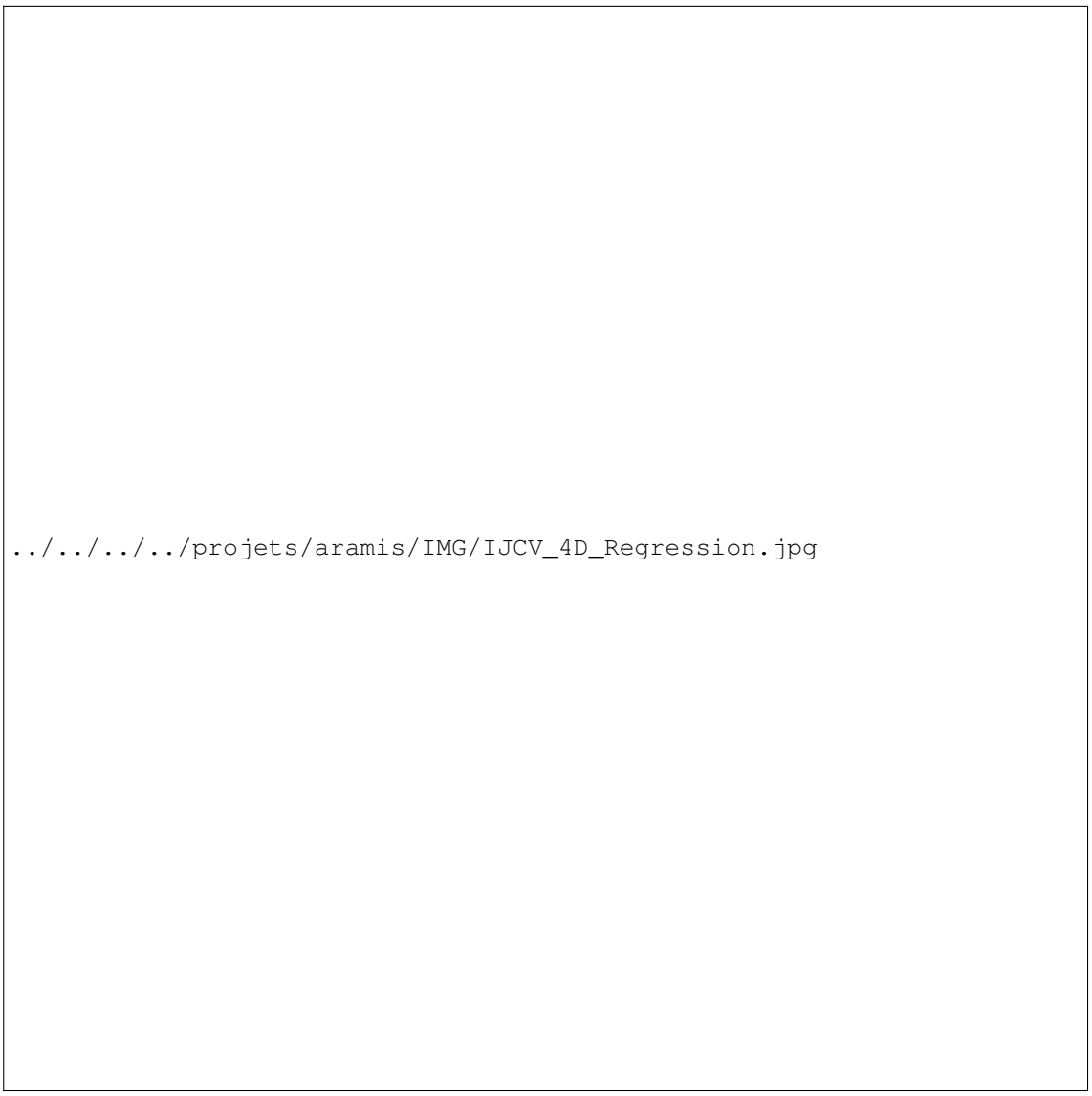
Shape regression is emerging as an important tool for the statistical analysis of time dependent shapes. We develop a new generative model which describes shape change over time, by extending simple linear regression to the space of shapes represented as currents in the large deformation diffeomorphic metric mapping (LDDMM) framework. By analogy with linear regression, we estimate a baseline shape (intercept) and initial momenta (slope) which fully parameterize the geodesic shape evolution. This is in contrast to previous shape regression methods which assume the baseline shape is fixed. We further leverage a control point formulation, which provides a discrete and low dimensional parameterization of large diffeomorphic transformations. This flexible system decouples the parameterization of deformations from the specific shape representation, allowing the user to define the dimensionality of the deformation parameters. We present an optimization scheme that estimates the baseline shape, location of the control points, and initial momenta simultaneously via a single gradient descent algorithm.

Shapes can be given as 3D meshes (as in [32]) or as 3D images (as in [31]).

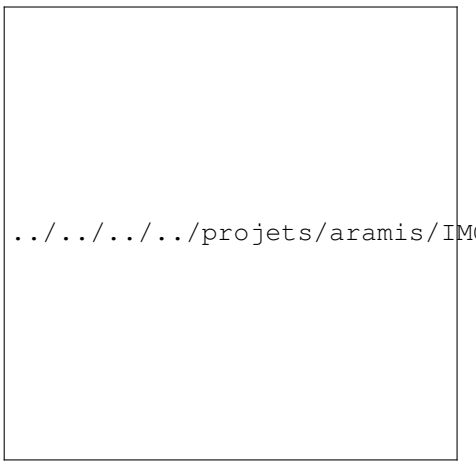
Related publications: [32], [31].

6.8. Discriminating brain microbleeds using phase contrast MRI in a multicentre clinical dataset

Participants: Takoua Kaaouana [Correspondant], Marie Chupin, Didier Dormont, Ludovic de Rochefort, Thomas Samaille.



a- Temporal regression of endocasts of bonobos (top) and chimpanzees (bottom)



b - morphological deformation

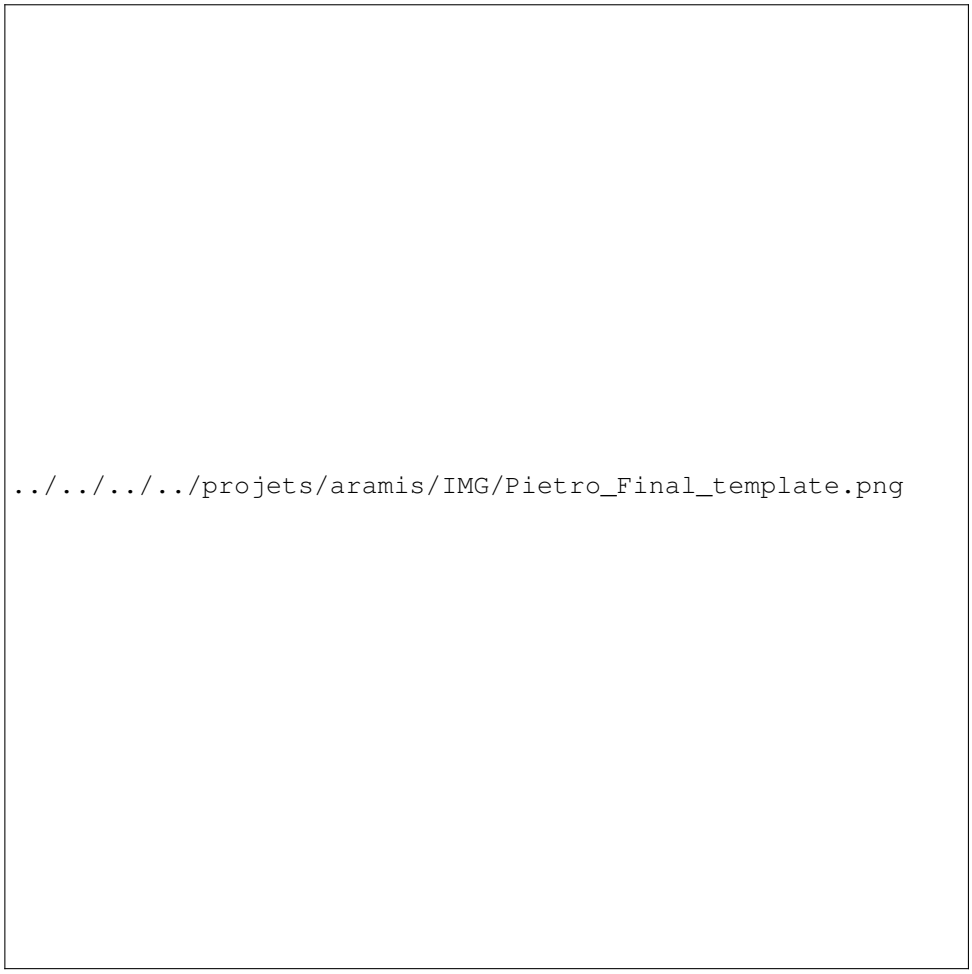


Figure 4. Atlas construction from a data set of left caudate nucleus and its associated fiber bundle that were segmented in images of patients with Gilles de la Tourette syndrome and controls. An initial template complex determines the topology of the model. Its shape is optimized given the patients data or the controls data only, thus resulting in two atlases showing different distributions of the fibers on the surface of the nucleus.

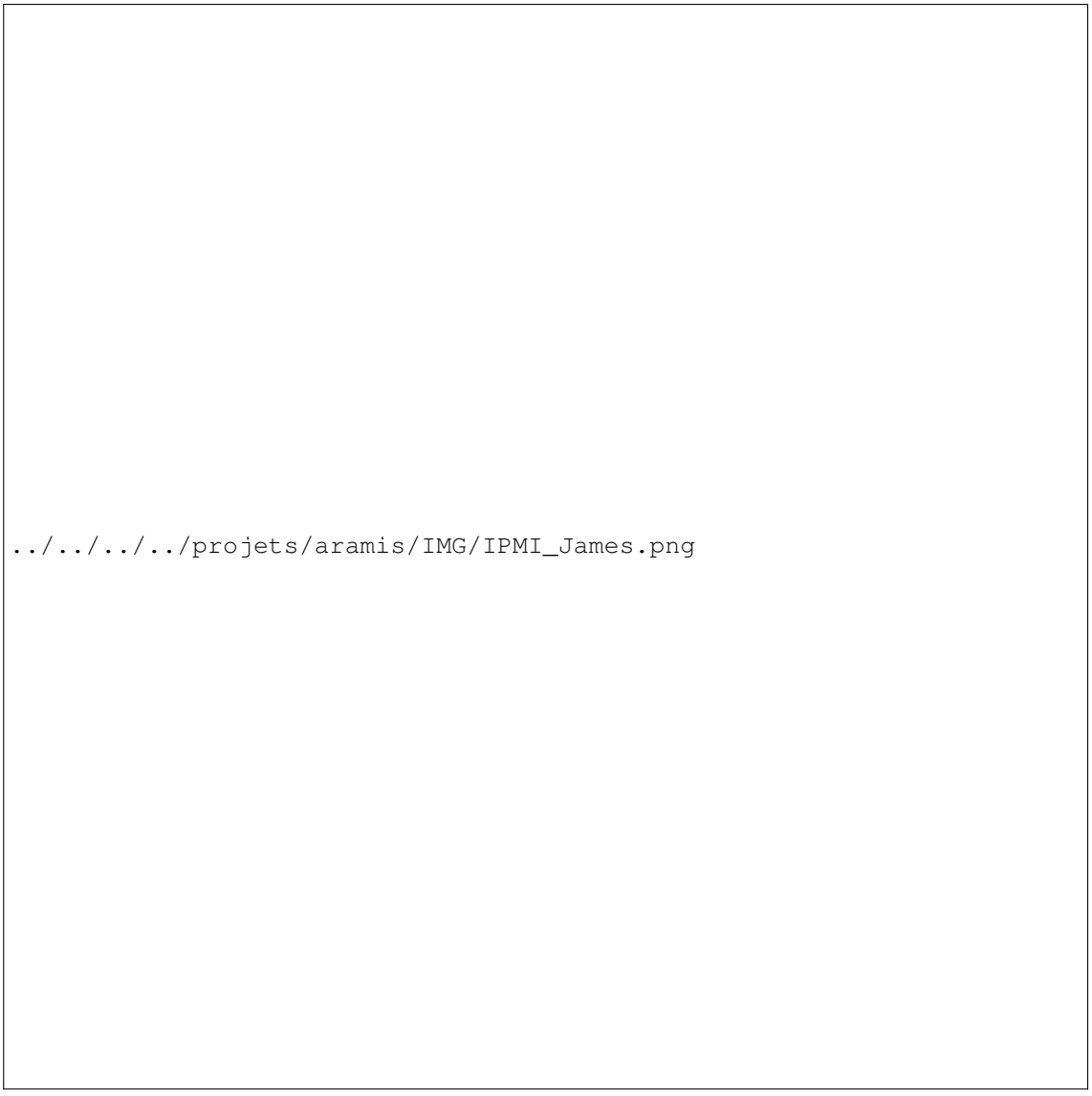


Figure 5. Average development of genu fiber tract from 2 to 24 months. Top row shows observed data for all subjects, which is clustered around 2, 12, and 24 months. Bottom row shows genu fiber tracts estimated from geodesic regression at several time points with velocity of fiber development displayed on the estimated fibers.

Brain microbleeds (BMBs) have emerged as a new imaging marker of small vessel diseases and they may play a crucial role in degenerative pathology such as Alzheimer's disease. Composed of hemosiderin, BMBs can be efficiently detected with MRI sequences sensitive to magnetic susceptibility (e.g. gradient recalled echo T2*W images). Nevertheless, that identification remains challenging because of confounding structures and lesions. Most T2*-weighted hyposignals result from local magnetic field inhomogeneity and can be identified either as BMBs, veins or brain micro-calcifications (BMCs). Differential diagnosis of BMBs and BMCs usually requires an additional CT scan. Quantitative susceptibility mapping techniques were proposed to discriminate between diamagnetic and paramagnetic structures, but they require a full 3D dataset and complex post-processing. We introduced a fast 2D phase processing technique including unwrapping and harmonic filtering thus yielding the internal field map, namely the field map generated only by sources within the volume of interest. We demonstrate its applicability and robustness on multicenter data acquired in standardized clinical settingand and its ability to discriminate between paramagnetic BMBs and diamagnetic BMCs through the use of the orientation of the dipolar pattern.

Related publications: [36].

6.9. Network symmetries and functional modules in the brain

Participants: Vincenzo Nicosia, Miguel Valencia, Mario Chavez [Correspondant], Albert Diaz-Guilera, Vito Latora.

We study the classical Kuramoto model in which the oscillators are associated to the nodes of a network and the interactions include a phase frustration, thus preventing full synchronization. The system organizes into a regime of remote synchronization where pairs of nodes with the same network symmetry are fully synchronized, despite their distance on the graph. We provide analytical arguments to explain this result and we show how the frustration parameter affects the distribution of phases. An application to brain networks suggests that anatomical symmetry plays a role in neural synchronization by determining correlated functional modules across distant locations.

Related publication: [19]

6.10. Accessibility of cortical networks during motor tasks

Participants: Mario Chavez [Correspondant], Fabrizio de Vico Fallani, Miguel Valencia, Mario Chavez, Julio Artieda, Vito Latora, Donatella Mattia, Fabio Babiloni.

Recent findings suggest that the preparation and execution of voluntary self-paced movements are accompanied by the coordination of the oscillatory activities of distributed brain regions. We used electroencephalographic source imaging methods to estimate the cortical movement-related oscillatory activity during finger ex- tension movements. We applied network theory to investigate changes (expressed as differences from the baseline) in the connectivity structure of cortical networks related to the preparation and execution of the movement. We computed the topological accessibility of different cortical areas, measuring how well an area can be reached by the rest of the network. Analysis of cortical networks revealed specific agglomerates of cortical sources that become less accessible during the preparation and the execution of the finger movements. The observed changes neither could be explained by other measures based on geodesics or on multiple paths, nor by power changes in the cortical oscillations.

Related publication: [3]

6.11. Abnormal functional connectivity between motor cortex and pedunculopontine nucleus following chronic dopamine depletion

Participants: Miguel Valencia, Mario Chavez [Correspondant], Julio Artieda, J. Paul Bolam, Juan Mena-Segovia.

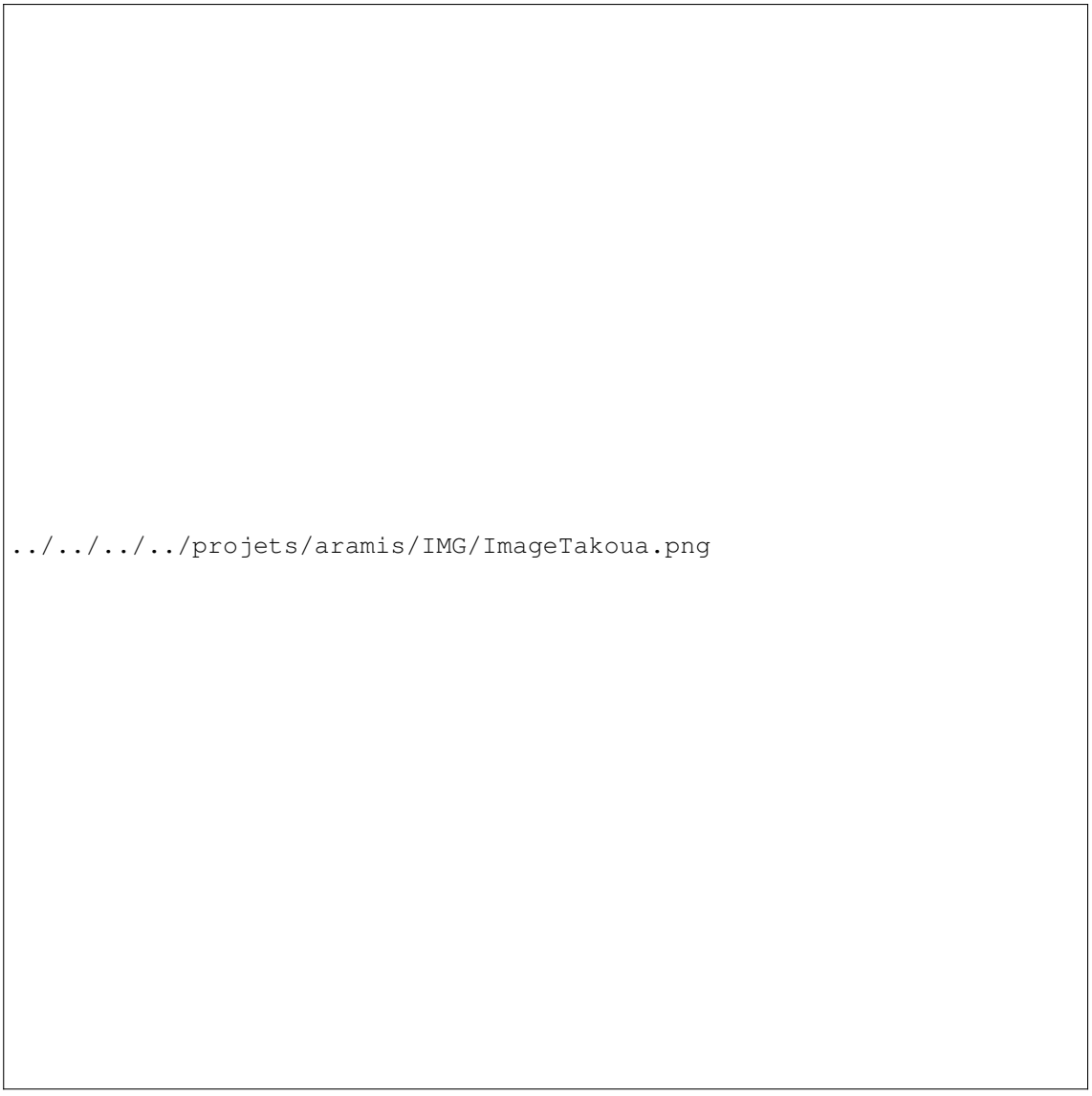


Figure 6. Philips (a,b,c,d) and Siemens (e,f,g,h) sample cases. Magnitude image (a,e), native phase image (b,f), and internal field map; axial (c,g) and sagittal (d,h). Zoom in white rectangle showing a dipolar pattern BMB (white arrow) and a physiologic calcification of the choroid plexus (black arrow).

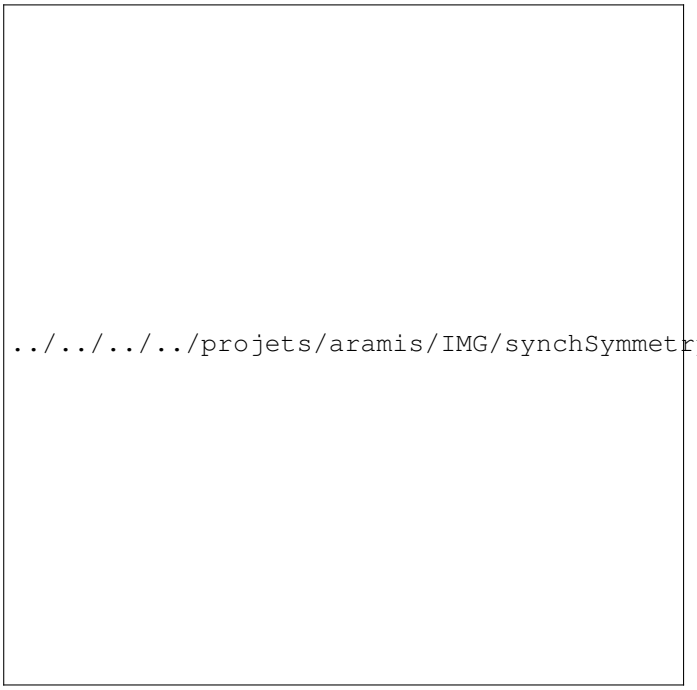


Figure 7. a) Brain areas with similar and dissimilar phases of the frustrated Kuramoto model are colored and superimposed onto an anatomical image. b) Examples of functional data from one subject recorded at the brain areas indicated in panel a). Colors are the same as those used in the anatomical image. c) Functional correlation (normalised values) Z between pairs of nodes as a function of their phase differences $\Delta\theta$ according to the simulated Kuramoto dynamics. The black solid curve corresponds to the average value over all the subjects, while the gray area covers the 5^{th} and the 95^{th} percentiles of the distribution. The dashed horizontal line indicates the threshold for statistical significant correlations ($p < 0.05$, corrected for multiple comparisons).

Figure 8. a) Averaged EMG and EEG (recorded at the postcentral region) signals of a subject during the execution of finger movements. Boxes define the three temporal epochs of EEG activity studied here: baseline (BASE), preparation (PRE) and execution period (EXE). Vertical dotted line indicates the movement onset. Examples of scalp and source-level networks obtained from one subject, at the frequency band Beta1, during the epoch EXE are shown in panels b) and c), respectively. Color map codes the number of connections.

The activity of the basal ganglia is altered in Parkinson's disease (PD) as a consequence of the degeneration of dopamine neurons in the substantia nigra pars compacta. This results in aberrant discharge patterns and expression of exaggerated oscillatory activity across the basal ganglia circuit. Altered activity has also been reported in some of the targets of the basal ganglia, including the pedunculopontine nucleus (PPN), possibly due to its close interconnectivity with most regions of the basal ganglia. However, the nature of the involvement of the PPN in the pathophysiology of PD has not been fully elucidated. We recorded local field potentials in the motor cortex and the PPN in the 6-hydroxydopamine (6-OHDA)-lesioned rat model of PD under urethane anesthesia. By means of linear and nonlinear statistics, we analyzed the synchrony between the motor cortex and the PPN, and the delay in the interaction between these two structures. We observed the presence of coherent activity between the cortex and the PPN in low- (5-15 Hz) and high-frequency bands (25-35 Hz) during episodes of cortical activation. In each case the cortex led the PPN. Dopamine depletion strengthened the interaction of the low-frequency activities by increasing the coherence specifically in the theta and alpha ranges and reduced the delay of the interaction in the gamma band. Our data show that cortical inputs play a determinant role in leading the coherent activity with the PPN, and support the involvement of the PPN in the pathophysiology of PD.

Related publication: [25]

6.12. Subthalamic Nucleus High-Frequency Stimulation Restores Altered Electrophysiological Properties of Cortical Neurons in Parkinsonian Rat

Participants: Bertrand Degos, Jean Michel Deniau, Mario Chavez [Correspondant], Nicolas Maurice.

Electrophysiological recordings performed in parkinsonian patients and animal models have confirmed the occurrence of alterations in firing rate and pattern of basal ganglia neurons, but the outcome of these changes in thalamo-cortical networks remains unclear. Using rats rendered parkinsonian, we investigated, at a cellular level in vivo, the electrophysiological changes induced in the pyramidal cells of the motor cortex by the dopaminergic transmission interruption and further characterized the impact of high-frequency electrical stimulation of the subthalamic nucleus, a procedure alleviating parkinsonian symptoms. We provided evidence that a lesion restricted to the substantia nigra pars compacta resulted in a marked increase in the mean firing rate and bursting pattern of pyramidal neurons of the motor cortex. These alterations were underlain by changes of the electrical membranes properties of pyramidal cells including depolarized resting membrane potential and increased input resistance. The modifications induced by the dopaminergic loss were more pronounced in cortico-striatal than in cortico-subthalamic neurons. Furthermore, subthalamic nucleus high- frequency stimulation applied at parameters alleviating parkinsonian signs regularized the firing pattern of pyramidal cells and restored their electrical membrane properties.

Related publication: [7]

6.13. Non-parametric resampling of random walks for spectral networks clustering

Participants: Fabrizio de Vico Fallani [Correspondant], Vincenzo Nicosia, Vito Latora, Mario Chavez.

Parametric resampling schemes have been recently introduced in complex network analysis with the aim of assessing the statistical significance of graph clustering and the robustness of community partitions. We proposed a method to replicate structural features of complex networks based on the non-parametric resampling of the transition matrix associated with an unbiased random walk on the graph. We tested this bootstrapping technique on synthetic and real-world modular networks and we showed that the ensemble of replicates obtained through resampling can be used to improve the performance of standard spectral algorithms for spectral clustering of graphs.

Related publication: [43]

6.14. Multiscale topological properties of functional brain networks during motor imagery after stroke

Participants: Fabrizio de Vico Fallani [Correspondant], Floriana Pichiorri, Giovani Morone, Marco Molinari, Fabio Babiloni, Febo Cincotti, Donatella Mattia.

In recent years, network analyses have been used to evaluate brain reorganization following stroke. However, many studies have often focused on single topological scales, leading to an incomplete model of how focal brain lesions affect multiple network properties simultaneously and how changes on smaller scales influence those on larger scales. In an EEG-based experiment on the performance of hand motor imagery (MI) in 20 patients with unilateral stroke, we observed that the anatomic lesion affects the functional brain network on multiple levels. In the beta (13–30 Hz) frequency band, the MI of the affected hand (Ahand) elicited a significantly lower smallworldness and local efficiency (Eloc) versus the unaffected hand (Uhand). Notably, the abnormal reduction in Eloc significantly depended on the increase in interhemispheric connectivity, which was in turn determined primarily by the rise of regional connectivity in the parieto-occipital sites of the affected hemisphere. Further, in contrast to the Uhand MI, in which significantly high connectivity was observed for the contralateral sensorimotor regions of the unaffected hemisphere, the regions with increased connectivity during the Ahand MI lay in the frontal and parietal regions of the contralaterally affected hemisphere. Finally, the overall sensorimotor function of our patients, as measured by Fugl–Meyer Assessment (FMA) index, was significantly predicted by the connectivity of their affected hemisphere. These results improve on our understanding of stroke-induced alterations in functional brain networks.

Related publication: [6]

6.15. Wavelet analysis in ecology and epidemiology: impact of statistical tests

Participants: Bernard Cazelles, Kevin Cazelles, Mario Chavez [Correspondant].

Wavelet analysis is now frequently used to extract information from ecological and epidemiological time series. Statistical hypothesis tests are conducted on associated wavelet quantities to assess the likelihood that they are due to a random process. Such random processes represent null models and are generally based on synthetic data that share some statistical characteristics with the original time series. This allows the comparison of null statistics with those obtained from original time series. When creating synthetic datasets, different techniques of resampling result in different characteristics shared by the synthetic time series. Therefore, it becomes crucial to consider the impact of the resampling method on the results. We have addressed this point by comparing seven different statistical testing methods applied with different real and simulated data. Our results showed that statistical assessment of periodic patterns is strongly affected by the choice of the resampling method, so two different resampling techniques could lead to two different conclusions about the same time series. Moreover, we showed the inadequacy of resampling series generated by white noise and red noise that are nevertheless the methods currently used in the wide majority of wavelets applications in epidemiology. Our results highlight that the characteristics of a time series, namely its Fourier spectrum and autocorrelation, are important to consider when choosing the resampling technique. Results suggest that data-driven resampling methods should be used such as the hidden Markov model algorithm and the ‘beta-surrogate’ method.

Related publication: [2]

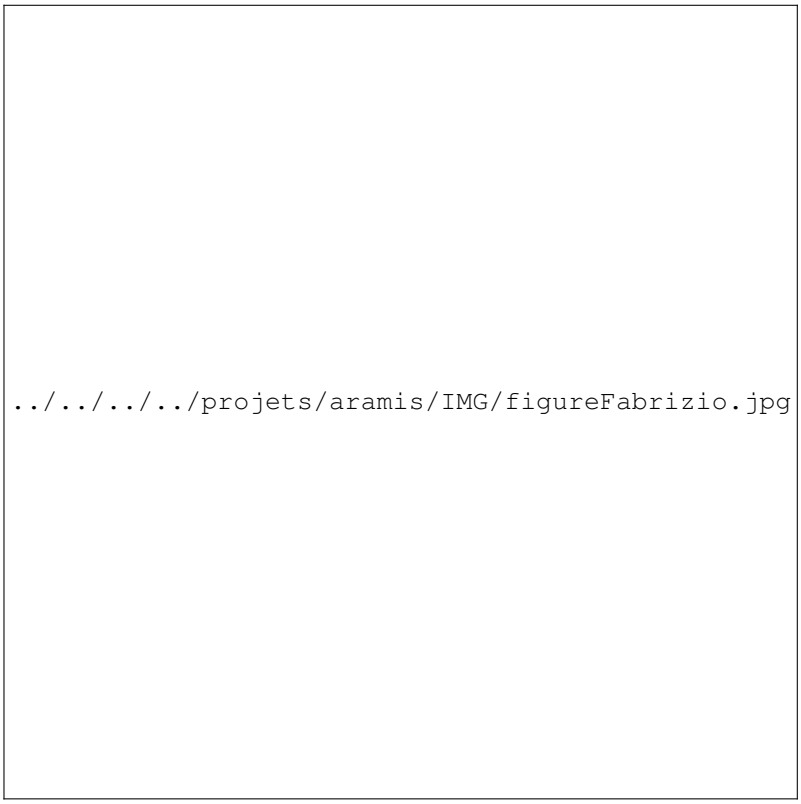


Figure 9. Grand average of brain networks in the Beta band during the MI of the unaffected Uhand and affected Ahand hand. Top plots: Scalp representation relative to Uhand (panel A) and Ahand (panel B) condition. Nodes are positioned according the actual EEG montage scheme. Blue and red lines denote the links within the unaffected (Uhemi) and the affected (Ahemi) hemisphere, respectively. Gray lines denote the inter-hemispheric links. The intensity of the color and the thickness of the lines vary as function of the number of patients exhibiting that significant link. Bottom part: graph representation of the brain networks relative to Uhand (panel A) and Ahand (panel B) condition. In this representation nodes are spatially repositioned through a force-based algorithm so that all the links are approximately of equal length with as few crossing edges as possible. Only links that were in common to more than 4 patients (20% of the sample) are illustrated here. Blue and red nodes indicate scalp electrodes placed over the undamaged (Uhemi) and damaged (Ahemi) hemisphere, respectively. The midline scalp electrodes (from Fpz to Oz) are illustrated as white nodes

ARLES Project-Team

6. New Results

6.1. Introduction

The ARLES project-team investigates solutions in the forms of languages, methods, tools and supporting middleware to assist the development of distributed software systems, with a special emphasis on mobile distributed systems enabling the ambient intelligence/pervasive computing vision.

Our research activities in 2013 have in particular accounted for the increasingly connected networking environment, as envisioned by the Future Internet, and further focused on one of its major components that is the Internet of Things, which allows connecting the physical with the digital word. In more detail, our research has focused on the following areas:

- Dynamic interoperability among networked systems toward making them eternal, by way of on-the-fly generation of connectors based on adequate system models (§ 6.2);
- Revisiting service-oriented computing toward the Future Internet, in particular dealing with the composition of highly heterogeneous services while ensuring quality of service (§ 6.3);
- Service oriented middleware for the ultra large scale future mobile Internet of Things (§ 6.4);
- Abstractions for enabling domain experts to easily compose applications on the Internet of Things (§ 6.5);
- Lightweight streaming middleware for the Internet of Things (§ 6.6); and
- Dynamic decision networks for decision-making in self-adaptive systems (§ 6.7).

6.2. Emergent Middleware

Participants: Emil Andriescu, Amel Bennaceur, Valérie Issarny.

Interoperability is a fundamental challenge for today's extreme distributed systems. Indeed, the high-level of heterogeneity in both the application layer and the underlying infrastructure, together with the conflicting assumptions that each system makes about its execution environment hinder the successful interoperation of independently developed systems. At the application layer, components may exhibit disparate data types and operations, and may have distinct business logics. At the middleware layer, they may rely on different communication standards, which define disparate data representation formats and induce different architectural constraints. Finally, at the network layer, data may be encapsulated differently according to the network technology in place.

A wide range of approaches have thus been proposed to address the interoperability challenge, as surveyed in [26]. However, solutions that require performing changes to the systems are usually not feasible since the systems to be integrated may be built by third parties (e.g., COTS —Commercial Off-The-Shelf— components or legacy systems); no more appropriate are approaches that prune the behavior leading to mismatches since they also restrict the systems' functionality. Therefore, many solutions that aggregate the disparate systems in a non-intrusive way have been investigated. These solutions use intermediary software entities, called *mediators*, to interconnect systems despite disparities in their data and/or interaction models by performing the necessary coordination and translations while keeping them loosely-coupled. However, creating mediators requires a substantial development effort and a thorough knowledge of the application-domain, which is best understood by domain experts. Moreover, the increasing complexity of today's distributed systems, sometimes referred to as Systems of Systems, makes it almost impossible to develop 'correct' mediators manually; correct mediators guarantee that the components interact without errors (e.g., deadlocks) and reach their termination successfully. Therefore, formal approaches are used to synthesize mediators automatically.

We posit that interoperability should neither be achieved by defining yet another middleware nor yet another ontology but rather by exploiting existing middleware together with knowledge encoded in existing domain ontologies to synthesize and implement mediators automatically. In [2], we have introduced the notion of *emergent middleware* for realizing mediators, which was initiated as part of the FP7 FET IP CONNECT project. Our work during the year 2013 has more specifically focused on the further elaboration of a comprehensive approach to mediator synthesis, including dealing with interoperability across protocol layers.

Mediator synthesis for emergent middleware: We focus on functionally-compatible components, i.e., components that at some high level of abstraction require and provide compatible functionalities, but are unable to interact successfully due to mismatching interfaces and behaviors. To address these differences without changing the components, mediators that systematically enforce interoperability between functionally-compatible components by mapping their interfaces and coordinating their behaviors are required. Our approach for the automated synthesis of mediators is performed in several steps.

The first step is interface matching, which identifies the semantic correspondence between the actions required by one component and those provided by the other. We incorporate the use of ontology reasoning within constraint solvers, by defining an encoding of the ontology relations using arithmetic operators supported by widespread solvers, and use it to perform interface matching efficiently. For each identified correspondence, we generate an associated matching process that performs the necessary translations between the actions of the two components' interfaces. The second step is the synthesis of correct-by-construction mediators. To do so, we analyze the behaviors of components so as to generate the mediator that combines the matching processes in a way that guarantees that the two components progress and reach their final states without errors. The synthesised mediator is the most general component that ensures freedom of both communication mismatches and deadlock in the composition of the components [15]. The last step consists in making the synthesized mediator concrete by incorporating all the details about the interaction of components. To do so, we compute the translation functions necessary to reconcile the differences in the syntax of the input/output data used by each component and coordinate the different interaction patterns that can be used by middleware solutions.

We refer the interested reader to [7] for a complete description of the approach. Our contribution primarily lies in handling interoperability from the application to the middleware layer in an integrated way. The mediators we synthesize act as: (i) translators by ensuring the meaningful exchange of information between components, (ii) controllers by coordinating the behaviors of the components to ensure the absence of errors in their interaction, and (iii) middleware by enabling the interaction of components across the network so that each component receives the data it expects at the right moment and in the right format.

Automated mediation for cross-layer protocol interoperability: Existing approaches to interoperability are restricted to solving either application heterogeneity when the underlying middlewares are compatible, or solving middleware heterogeneity at each protocol layer separately. In real world scenarios, this does not suffice: application and middleware boundaries are ill-defined and solutions to interoperability must consider them in conjunction. We have been studying the case of cross-layer interoperability where protocol mediation is performed between protocol stacks, rather than between protocol layers separately. Such interoperability approaches are appropriate for systems that rely on complex protocol stacks, where application and middleware layers are tightly coupled.

Systems relying on tightly coupled protocol stacks exchange complex messages that consist of a composition of heterogeneous data formats. To enable interoperation, complex messages from one system must be translated into a different complex format that another system accepts such that the two can interact. While Off-The-Shelf and third party message parsers are widely available for simple message formats (i.e., message formats corresponding to a single protocol layer), complex message formats are typically unique since they are the result of a protocol binding. Protocol binding represents the connection between one protocol and another to create a new communication flow. Some middleware protocols recommend or restrict to certain types of default binding (e.g., HTTP provides an extensive set of rules for binding, such as Content-Encoding and Content-Type). However, real systems are often designed following a custom binding mechanism, restricting the application of automated mediation solutions. This problem occurs primarily because complex message formats cannot be easily interpreted.

Many solutions address this composition issue by introducing Domain Specific Languages that can be used by experts to specify parsers for complex message formats. Yet, whenever messages have a more complicated syntax, providing their DSL descriptions becomes difficult as well. Further, such approaches are not future proof as more protocols are expected to emerge, which will not be accounted for by DSLs that are defined according to known message formats. An alternative is to generate parsers based on the composition of third-party parsers that are usually included with protocol implementations. However, third-party parsers cannot be used unless the protocol binding rules are identified by an expert, further allowing to implement the bridge between one parser's output data and the other parser's input data. To this end, we designed an approach for generating composed parsers that can process complex messages, accompanied by a formal mechanism for defining complex message formats based on existing data formats. Our approach relies on user-provided parser composition rules, which reflect the binding requirements of complex message formats.

We posit that our method is more efficient than implementing complex parsers, defining them using DSLs, or directly implementing the binding of protocols. Furthermore, with this solution, we support the automated synthesis of mediators at the application layer using the mapping-based approach discussed above, by automatically generating an abstract representation of the application data exchanged by the interoperating components.

6.3. Service-oriented Computing in the Future Internet

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Valérie Issarny, Ajay Kattepur.

With an increasing number of services and devices interacting in a decentralized manner, *choreographies* represent a scalable framework for the Future Internet. The service oriented architecture inherent to choreographies allows abstracting multiple devices as components, that interact through middleware connectors via standard protocols. However, the heterogeneous nature of devices leads to choreographies that not only include conventional services, but also sensor-actuator networks, databases and service feeds. We reason about their behavior through abstract middleware interaction paradigms, such as client-service (CS), publish-subscribe (PS) and tuple space (TS), made interoperable through the *eXtensible Service Bus* (XSB) connector.

Extensible Service Bus for the Future Internet: XSB is an abstract service bus that deals effectively with the cross-integration of heterogeneous interaction paradigms [17]. Inside the XSB, the CS, PS and TS paradigms are modeled as abstract base connectors. Their *space coupling* semantics are represented with programming interfaces used by applications (APIs) and corresponding application interface description languages (IDLs). Their behavioral semantics are formally specified in terms of LTS (Labeled Transition Systems). We formally verify the correctness of these behavioral specifications with respect to *time coupling* and *concurrency* properties expressed in LTL temporal logic. This allows stating the correctness of the connector models with respect to the semantics that they must have. This further enables identifying the behavioral semantics of the XSB connector derived from the interconnection of base connectors. More specifically, in order to identify the time coupling and concurrency semantics of XSB and construct a converter among the base connectors, we build upon the formal method of *protocol conversion via projections*[10]. According to this method, conversion between two different protocols is possible if both protocols can be projected (where projection is an abstraction defined as a set of transformations on the protocol LTS) to a *functionally sufficient* common *image protocol*. Then, the end-to-end protocol of the interconnection of the two protocols is this image protocol.

We have implemented our XSB solution into an extensible development and execution platform for application and middleware designers. Using this platform, they can easily develop composite applications: they only need to build descriptions for the constituent services and directives for data mapping among them. Our platform then deals with reconciling among the heterogeneous interaction paradigms and protocols of the services by employing *binding components (BCs)* that adapt between the native middleware of the services and the XSB bus protocol. The XSB itself is implemented on top of an existing ESB substrate. Support for new middleware platforms, new ESB substrates, or even new interaction paradigms can be incorporated in a facilitated way thanks to the provided XSB architectural framework.

[10] Lam, S.S.: Protocol Conversion. IEEE Trans. Softw. Eng. 14(3) (1988) 353–362.

QoS composition and analysis of heterogeneous choreographies: Leveraging on the functional interoperability across interaction paradigms offered by the XSB, we study the Quality of Service (QoS) performance of choreographies [21]. QoS dependency plays an important role in the service oriented system lifecycle, including discovery, runtime selection, replacement and contractual guarantees. Consequently, QoS composition among choreographed devices should tackle multi-dimensional probabilistic metrics combined with message passing constraints imposed at design-time. We make use of an algebraic QoS composition model that is applied at the interaction paradigm level to study the composition of QoS metrics, and the subsequent tradeoffs. While traditional QoS composition analysis has been done purely at the application level, analyzing the effect of middleware interactions allows us to study CS, PS and TS based device compositions. This produces interesting insights such as selection of a particular system and its middleware during design-time, or end-to-end QoS expectation/guarantees during runtime. Our formulation also allows for runtime reconfiguration, in order to optimally produce design time QoS expectations. Such flexible reconfiguration policies are crucial in the case of large scale choreographies with high variability in runtime performance of participating devices.

Further, we study the effect of time/space coupling on the latency of successful transactions across the XSB connector [20]. XSB models the message passing among peers through generic `post` and `get` operations, that represent peer behavior with both tight (CS) and loose (PS/TS) time/space coupling. The heterogeneous *lease* and *timeout* behaviors of these operations severely affect latency and success rates of messages passed either synchronously or through callbacks. By precisely studying the timing thresholds using timed automata models, we verify conditions for accurate message transactions with XSB connectors. This offers choreography designers the ability to set these timing thresholds (bottom-up) or select a particular interaction paradigm (top-down) for runtime enaction.

6.4. Service-oriented Middleware for the Mobile Internet of Things

Participants: Sara Hachem, Valérie Issarny, Georgios Mathioudakis, Animesh Pathak.

The Internet of Things (IoT) is characterized by an increasing number of Things embedding sensing, actuating, processing, and communication capacities. A considerable portion of those Things will be *mobile* Things, which come with several advantages yet lead to unprecedented challenges. The most critical challenges, that are directly inherited from, yet amplify, today's Internet issues, lie in handling i) the large scale of users and mobile Things, ii) providing interoperability across the heterogeneous Things, and iii) overcoming the unknown dynamic nature of the environment, due to the mobility of an ultra-large number of Things.

Service-Oriented Architecture (SOA) provides solid basis to address the above challenges as it allows the functionalities of sensors/actuators embedded in Things to be provided as services, while ensuring loose-coupling between those services and their hosts, thus abstracting their heterogeneous nature. In spite of its benefits, SOA has not been designed to address the ultra-large scale of the mobile IoT. Consequently, an alternative is provided within a novel Thing-based Service-Oriented Architecture, that revisits SOA interactions and functionalities, service discovery and composition in particular. The novel architecture is concretized within MobIoT, a middleware solution that is specifically designed to manage and control the ultra-large number of mobile Things in partaking in IoT-related tasks.

In accordance with SOA, MobIoT comprises *Discovery*, *Composition & Estimation*, and *Access* components, yet modifies their internal functionalities. In more detail, the Discovery component enables Thing-based service registration (for Things to advertise hosted services) and look-up (for Things to retrieve remote services of interest). In order to handle the ultra large number of mobile Things and their services in the IoT, the component revisits the Service-Oriented discovery and introduces *probabilistic discovery* to provide, not *all*, but only a sufficient *subset of services that can best approximate* the result that is being sought after [18], [11]. Furthermore, the Composition & Estimation component (C&E) provides automatic composition of Thing-based services. This capacity is of interest in the case where no service can perform a required measurement/action task directly (based on its atomic functionalities). Thing-based service composition executes in three phases: i) *expansion*, where composition specifications are automatically identified; ii) *mapping*, where actual service instances (running services) are selected based on their functionalities and the physical attributes of their hosts; and iii) *execution*, where the services are accessed and the composition specifications are executed.

Thing-based service composition revisits Service-Oriented composition by executing seamlessly with no involvement from developers or end users. Last but not least, the Access component provides an easy to use interface for developers to sample sensors/actuators while abstracting sensor/actuator hardware specifications. Additionally, it revisits Service-Oriented access by executing access to services transparently and wrapping access functionalities internally. Thus, it alleviates that burden from users, initially in charge of this task. The Access component supports access to remote services and to locally hosted services.

6.5. Composing Applications in the Internet of Things

Participants: Aness Bajia, Pankesh Patel, Animesh Pathak, Françoise Sailhan.

As introduced above, the Internet of Things integrates the physical world with the existing Internet, and is rapidly gaining popularity, thanks to the increased adoption of smart phones and sensing devices. Several IoT applications have been reported in recent research, and we expect to see increased adoption of IoT concepts in the fields of personal health, inventory management, and domestic energy usage monitoring, among others.

An important challenge to be addressed in the domain of IoT is to enable domain experts (health-care professionals, architects, city planners, etc.) to develop applications in their fields rapidly, with minimal support from skilled computer science professionals. An ideal application development abstraction of the IoT will allow (domain expert) developers to intuitively specify the rich interactions between the extremely large number of disparate devices in the future Internet of Things. The goal of our research is then to propose a suitable application development framework, where our work this year covered the two following related areas.

Multi-stage model-driven approach for IoT application development: We have proposed a multi-stage model-driven approach for IoT application development based on a precise definition of the role to be played by each stakeholder involved in the process: domain expert, application designer, application developer, device developer, and network manager [22]. The metamodels/abstractions available to each stakeholder are further customized using the inputs provided in the earlier stages by other stakeholders. We have also implemented code-generation and task-mapping techniques to support our approach. Our evaluation based on two realistic scenarios shows that the use of our techniques/framework succeeds in improving productivity in the IoT application development process. More details of our approach can be found in [8].

Integrating support for non-functional requirements while programming IoT applications: Given that devices and networks constituting the IoT are prone to failure and consequent loss of performance, it is natural that IoT applications are expected to encounter and tolerate several classes of faults - something that still largely remains within the purview of low-level-protocol designers. As part of our work on the MURPHY project (§ 7.1.1.1), we are addressing this issue by proposing: i) a set of abstractions that can be used during macroprogramming to express fault tolerance requirements, and ii) a runtime system that employs adaptive fault tolerance (AFT) to provide fault tolerance to the sensing application. Complementary to this, we have proposed task mapping algorithms to satisfy those requirements through a constraint programming approach [19]. Through evaluations on realistic application task graphs, we show that our constraint programming model can effectively capture the end-to-end requirements and efficiently solves the combinatorial problem introduced.

We have continually incorporating our research results in the above areas into *Srijan* (§ 5.6), which provides an easy-to-use graphical front-end to the various steps involved in developing an application using the ATaG macroprogramming framework.

6.6. Lightweight Streaming Middleware for the Internet of Things

Participants: Benjamin Billet, Valérie Issarny.

The Internet of Things (IoT) is a promising concept toward pervasive computing as it may radically change the way people interact with the physical world. One of the challenges raised by the IoT is the in-network continuous processing of data streams presented by Things, which must be investigated urgently because it affects the future data models of the IoT. This cross-cutting concern has been previously studied in the context of Wireless Sensor and Actuator Networks (WSAN) given the focus on the acquisition and in-network processing of sensed data. However, proposed solutions feature heterogeneous technologies that are difficult to integrate and complex to use, which represents a hurdle to their wide deployment. In addition, new types of smart sensors are emerging due to technological advances (e.g., Oracle SunSpot), enabling the implementation of complex processing tasks directly into the network, without using proxies or sending every data to the cloud. There is thus a need for a distributed middleware solution for data stream management that leverages existing WSAN work, while integrating it with today's Web technologies in order to improve the flexibility and the interoperability of the future IoT. Toward that goal, we have been developing Dioptase, a Data Stream Management System for the IoT, which aims to integrate the Things and their streams into today's Web by presenting sensors and actuators as services. The middleware specifically provides a way to describe complex fully-distributed stream-based mashups and to deploy them dynamically, at any time, as task graphs, over available Things of the network, including resource-constrained ones. To this end, Dioptase enables task graphs to be composed of Thing-specific tasks (directly implemented on the Thing) and dynamic tasks that communicates using data streams. Dynamic tasks are then described in a lightweight DSL, which is directly interpreted by the middleware and provides specific primitives to manipulate data streams.

As part of the design of Dioptase, we have been investigating dedicated task mapping. Task mapping, which basically consists of mapping a set of tasks onto a set of nodes, is a well-known problem in distributed computing research. However, as a particular case of distributed systems, the Internet of Things (IoT) poses a set of renewed challenges, because of its scale, heterogeneity and properties traditionally associated with WSAN, shared sensing, continous processing of data streams and real time computing. To handle IoT features, we present a formalization of the task mapping problem that captures the varying consumption of resources and various constraints (location, capabilities, QoS) in order to compute a mapping that guarantees the lifetime of the concurrent tasks inside the network and the fair allocation of tasks among the nodes (load balancing). It results in a binary programming problem for which we provide an efficient heuristic that allows its resolution in polynomial time. Our experiments show that our heuristic: (i) gives solutions that are close to optimal and (ii) can be implemented on reasonably powerful Things and performed directly within the network, without requiring any centralized infrastructure.

6.7. Dynamic Decision Networks for Self-Adaptive Systems

Participants: Amel Belaggoun, Nelly Bencomo, Valérie Issarny, Peter Sawyer.

Different modeling techniques have been used to model requirements and decision-making of self-adaptive systems [25]. Important successful techniques based on goal models have been prolific in supporting decision-making according to partial and total fulfillment of functional (goals) and non-functional requirements (softgoals). The final decision about what strategy to use is based on a utility function that takes into account the weighted sum of the different effects of the non-functional requirements. Such solutions have been used both at design and run time including our own solutions using runtime goal models. Different modeling techniques have been used to model requirements and decision-making of self-adaptive systems [25]. Important successful techniques based on goal models have been prolific in supporting decision-making according to partial and total fulfillment of functional (goals) and non-functional requirements (softgoals). The final decision about what strategy to use is based on a utility function that takes into account the weighted sum of the different effects of the non-functional requirements. Such solutions have been used both at design- and run-time including our own solutions using runtime goal models.

We have enriched the decision-making supported by goal models with the use of Bayesian Dynamic Decision Networks (DDNs) [12]. Our novel approach supports reasoning about partial satisfaction of soft-goals using probabilities and uses machine learning. When using DDNs, we introduce new ways to tackle uncertainty based on probabilities that can be updated based on runtime evidence. We have reported the results of the

application of the approach on two different cases, one of them being the case of dynamic reconfiguration of a remote data mirroring network that must spread data among servers while minimizing costs and loss of data. Our early results suggest the decision-making process of self-adaptive systems can be improved by using DDNs.

This work has been developed under the umbrella of the Marie Curie Project Requirements@run.time (§ 7.2.1.4). The main results achieved during the year 2013 are:

- A Bayesian-based technique to support the decision making of self-adaptive systems [14]. DDN-based approaches adopt probabilistic methods (i.e., Bayesian methods) and decision theory to assess the consequences of uncertainty. Using the approach, suitable choices to satisfice functional requirements of the system are identified from a range of alternative decisions and their expected utilities. Satisficement of NFRs is modeled using conditional probabilities given the design decisions. Preferences over decisions are modeled using weights associated with pairs of design alternatives and NFRs, and used when computing the expected utilities of the architectural design alternatives. The decision taken by the DDN is that with the highest expected utility. The approach offers the benefits of machine learning.
- A formal Bayesian definition of surprise as the basis for quantitative analysis to measure degrees of uncertainty and deviation of self-adaptive systems from normal behavior [13]. Specifically, a Bayesian surprise quantifies how new evidence affects assumptions of the world (properties in the models). A “surprising" event may provoke a large divergence between the beliefs distributions prior and posterior to that event. As such and depending on how big or small this divergence is, the running system may decide to either: (i) dynamically adapt accordingly, or (ii) temporarily avoid any action of adaptation and flag up the fact that a potential abnormal situation has been found. While doing (ii) we are offering a specific implementation of the RELAX language previously developed by Bencomo and her co-authors.

AXIS Project-Team

6. New Results

6.1. Introduction

Our new results are split into our three sub-objectives as described in Section 3.1 :

- **Sub-Objective 1: Mining for Knowledge Discovery in Information Systems:**
  This year we obtained ten main results (cf. Section 6.2): five on Clustering methods, four on how to apply these clustering methods on real data and finally one related to the use of ontology for Multi-View KDD process.

  Let us note that two 2011 results have been published this year as book chapters [34], [31].
  Chongsheng Zhang published also his work conducted during his Explore programm at UCLA (USA) when, as AxIS PhD student, he was visiting the WIS team of Prof. Carlo Zaniolo at UCLA in 2010 [26].
- **Sub-Objective 2: Information and Social Networks Mining for Supporting Information Retrieval:**
  This year, we pursued our two main works on this topic (cf. Section 6.3):
  - the detection of communities in a social network (detection of graphs extracted from relational data) (cf. Section 6.3.1),
  - the multi view clustering of relational data (cf. SEction 6.3.2).
- **Sub-Objective 3: Interdisciplinary Research For Supporting User Oriented Innovation:**
  With the expansion of the innovation community beyond the firm's boundaries (the so-called "open innovation") a lot of changes have been introduced in design and evaluation processes: the users can become co-designers, HCI design and evaluation focus is no longer placed on usability only but also on the whole user experience [70] [11] , experimentations take place out of labs with large numbers of heterogeneous people instead of carefully controlled panels of users etc.

  All these deep changes required improvements of existing practices, methods and tools for the design/evaluation of information systems as well as for usage analysis. This evolution called also for a structured user-centred methodology (methods and ICT tools) to deal with open innovation. Various different disciplines and trends are dedicated in understanding user behaviour on Internet and with Digital Technologies, notably Human Computer Interaction community (HCI), Computer Supported Cooperative Work (CSCW), Workplace Studies, Service Design, Distributed Cognition and Data Mining.

  Our contribution to open innovation research related to ICT-based services or products keeps its focus on usage analysis and user experience measurement for design, evaluation and maintenance of information systems and our activities from 2011 have been conducted both breadth wise and in depth with two main objectives :
  - Improving design and evaluation support tools and methods for user driven driven innovation,
  - Development of the FocusLab platform

This year, our research was conducted along three focus:

- Extension of usability methods and models (cf. Section 6.4). First we pursued our work on User Evaluation and Tailoring of Personal Information in the context of the ANR project PIMI. Second a paper related to our strategy and heuristics for rural tourist web sites benchmarking elaborated in the context of the Pacalabs project HOTEL-REF-PACA is written for submission in 2014;
- Designing and evaluating user experience in the context of a living lab: this year five results came from ELLIOT project (cf. Section 6.5) such as an environmental data platform based on citizen sensing, low-cost sensor, user experience measurement, user behaviour change analysis, studies of persuasive technologies and gamification in Energy economy and green services.
- FocusLab Platform (cf. Section 6.6).

6.2. Mining for Knowledge Discovery in Information Systems

6.2.1. Fuzzy Clustering on Multiple Dissimilarity Matrices

Participants: Yves Lechevallier, Francisco de Carvalho.

During 2013 we introduce fuzzy clustering algorithms [18] and [27] that can partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. The aim is to obtain a collaborative role of the different dissimilarity matrices to get a final consensus partition. These matrices can be obtained using different sets of variables and dissimilarity functions. These algorithms are designed to furnish a partition and a prototype for each fuzzy cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fit between the fuzzy clusters and their representatives. These relevance weights change at each algorithm iteration and can either be the same for all fuzzy clusters or different from one fuzzy cluster to another.

A new algorithm [19] based on a non-linear aggregation criterion, weighted Tchebycheff distances, more appropriate than linear combinations (such as weighted averages) for the construction of compromise solutions is proposed.

Experiments with real-valued data sets from the UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/) as well as with interval-valued and histogram-valued data sets show the usefulness of the proposed fuzzy clustering algorithms.

6.2.2. Clustering of Functional Boxplots for Multiple Streaming Time Series

Participant: Yves Lechevallier.

We introduced a micro-clustering strategy for Functional Boxplots [30]. The aim is to summarize a set of streaming time series split in non overlapping windows. It is a two step strategy which performs at first, an on-line summarization by means of functional data structures, named Functional Boxplot micro-clusters; then it reveals the final summarization by processing, off-line, the functional data structures. Our main contribution consists in providing a new definition of micro-cluster based on Functional Boxplots and, in defining a proximity measure which allows us to compare and update them. This allows us to get a finer graphical summarization of the streaming time series by five functional basic statistics of data. The obtained synthesis will be able to keep track of the dynamic evolution of the multiple streams.

This work is done in collaboration with the laboratory of Political Science "Jean Monnet", Second University of Naples, Caserta, Italy.

6.2.3. Web Page Clustering based on a Community Detection Algorithm

Participant: Yves Lechevallier.

Extracting knowledge from Web user's access data in Web Usage Mining (WUM) process is a challenging task that is continuing to gain importance as the size of the Web and its user-base increase. That is why meaningful methods have been proposed in the literature in order to understand the behaviour of the user in the Web and improve the access modes to information.

During 2013 we pursued our previous work on our approach for extracting data based on the modularity function. This approach discovers the existing communities by modeling the data obtained in the pre-processing operation as a weighted graph. The method discriminates the communities through their subject of interest and extract relevant knowledge.

This work is done in collaboration with Yacine Slimani from the LRIA laboratory at the Ferhat Abbas University, Setif, Algerie and will be submitted to an international journal.

6.2.4. Normalizing Constrained Symbolic Data for Clustering

Participants: Marc Csernel, Francisco de Carvalho.

Clustering is one of the most common operation in data analysis while constrained is not so common. During 2013 we presented a clustering method [31] in the framework of Symbolic Data Analysis (S.D.A) which allows us to cluster Symbolic Data. Such data can be constrained relations between the variables, expressed by rules which express the domain knowledge. But such rules can induce a combinatorial increase of the computation time according to the number of rules. The algoritm presented a way to cluster such data in polynomial time. This method is based first on the decomposition of the data according to the rules, then we can apply to the data a clustering algorithm based on dissimilarities.

6.2.5. Dynamic Clustering Method for Mixed Data

Participants: Yves Lechevallier, Marc Csernel, Brigitte Trousse.

For ELLIOT project purposes (cf. Section 7.3.1), a new version of MND method (Dynamic Clustering Method for Mixed Data) has been elaborated. It determines iteratively a series of partitions which improves at each step the underlying clustering criterion. All the proposed distance functions for p variables are determined by sums of dissimilarities corresponding to the univariate component descriptors Y_j. The most appropriate dissimilarities have been suggested above according to the type of variables.

In practice, however, data to be clustered are typically described by different types of variables. An overall dissimilarity measure is obtained by a linear combination of the dissimilarity measures computed with respect to the different kinds of variables.

A new release of MND algorithm based on past work [80] has been developed for ELLIOT purposes, providing some default configuration parameters for non experts.

In this version two types of distances are proposed:

- **Quantitative distance**: the choice is type L1 distance or Euclidean distances when the types of variables are quantitative or continuous.
- **Boolean distance**: the choice is Khi2, type L1 distance or Euclidean distances when the type of variables is categorical or discrete.

This algorithm has been applied to cluster answers at questionnaires issued from a diary tool within the ELLIOT Green Services use case (cf. Section 6.5.4).

6.2.6. Applying a K-means clustering method for districts clustering according to Pollution

Participants: Brigitte Trousse, Yves Lechevallier, Guillaume Pilot, Caroline Tiffon.

Our motivation was to provide citizen a comparative analysis at the district level related to pollution data from Azimut stations (ozone O3 and nitrogen dioxide NO2). To achieve this, the Nice Côte d'Azur territory was discretized into small areas. IoT Data are preprocessed for each district and period of time before applying clustering. The temporal and spatial units were clustered into 5 and then into 6 clusters. The partition into 5 clusters was selected, then the temporal units for each area were counted. For the partition in 5 clusters, for each area the percent of each cluster was counted. Around 30 areas with more than 10 temporal units were found. We improved this to classify different districts of the city based on their IoT data (Azimut data O3-NO2) for each hour/day in order to provide a new functionnality in the second version of MyGreenServices.

This work is partially funded by ELLIOT project (see Section 7.3.1).

6.2.7. *Summarizing Dust Station IoT Data with REGLO, a FocusLab web service*

Participants: Yves Lechevallier, Brigitte Trousse, Guillaume Pilot, Xavier Augros.

Within ELLIOT, we applied the GEAR (or REGLO in French) method [57], [58], [59] on the evolution of dust data issued from one citizen sensor.

Our motivation was to summarize IoT data in order to have a pollution context for each user. Such IoT summaries constitute interesting individual contextual data for supporting the living lab manager to better interpret the user behavior and finally the user experience.

REGLO summarised IoT data with isolated points and line segments.

The goal now is to carry out an analysis of these summaries to automatically determine the characteristics of the curve.

We selected only segments. For each segment we calculated four variables that characterize it:

- The slope of the segment,
- The midpoint of the segment (average of this segment),
- The length of the segment,
- The duration of the segment (the time interval between the start time and the end time of the segment).

From these four values we can achieve an interpretation of the previous curve, taking into account only two variables and constructing a 2D representation.

This work is partially funded by ELLIOT project (see Section 7.3.1).

6.2.8. *Clustering of Solar Irradiance*

Participants: Thierry Despeyroux, Francisco de Carvalho, Yves Lechevallier, Thien Phuc Hoang Nguyen.

The development of grid-connected photovoltaic power systems leads to new challenges. The short or medium term prediction of the solar irradiance is definitively a solution to reduce the storage capacities and, as a result, authorizes to increase the penetration of the photovoltaic units on the power grid. We present the first results of an interdisciplinary research project which involves researchers in energy, meteorology and data mining, addressing this real-world problem. The objective here is to show interest and disadvantages of two approaches for classifying curves.

In Reunion Island from December 2008 to March 2012, solar radiation measurements has been collected, every minutes, using calibrated instruments. Prior to prediction modelling, two clustering strategies has been applied for analysis the data base of 951 days.

During 2013 we continued our research and obtained many results [28].

Our methodology is based on two clustering approaches. The objective here is to show interest and disadvantages of two approaches for classifying curves.

The first approach combines the following proven data-mining methods. Principal Component Analysis was used as a pre-process for reduction and de-noising and the Ward Hierarchical and K-means methods to find a partition with a good number of classes.

The second approach [78],[20] uses a clustering method that operates on a set of dissimilarity matrices. Each cluster is represented by an element or a subset of the set of objects to be classified. The five meaningfully clusters found by the two clustering approaches are compared.

6.2.9. Understanding of Cooking User's Recipes by Extracting Intrinsic Knowledge

Participants: Damien Leprovost, Thierry Despeyroux, Yves Lechevallier.

On community web sites, users share knowledge, being both authors and readers. We present a method to build our own understanding of the semantics of the community, without the use of any external knowledge base. We perform this understanding by knowledge extraction from analysed user contributions. We propose an evaluation of the trust attributable to that deduced understanding to assess the quality of user content, on cooking recipes provided by users on sharing web sites. This work is partially funded by FIORA project (see Section 7.2.2). Two articles have been accepted in early 2014 [25], [29].

6.2.10. Knowledge Modeling for Multi-View KDD Process

Participant: Brigitte Trousse.

We pursued our supervision (with our colleagues H. Behja and A. Marzark from Morocco) of E.L. Moukhtar Zemmouri's PhD thesis (Morocco) on a Viewpoint Model in the context of a KDD process, topic we initiated during Behja's PhD thesis [40]). E. Zemmouri defended his thesis at the end of this year [75]. Below is the summary of his PhD thesis.

Knowledge Discovery in Databases (KDD) is a highly complex, iterative and interactive process aimed at the extraction of previously unknown, potentially useful, and ultimately understandable patterns from data. In practice, a KDD process involves several actors (domain experts, data analysts, KDD experts â€¦) each with a particular viewpoint. We define a multi-view analysis as a KDD process held by several experts who analyze the same data with different viewpoints. We propose to support users of multi-view analysis through the development of a set of semantic models to manage knowledge involved during such analysis. Our objective is to enhance both the reusability of the process and coordination between users. To do so, we propose first a formalization of Viewpoint in KDD and a Knowledge Model that is a specification of the information and knowledge structures and functions involved during a multi-view analysis. Our formalization, using OWL ontologies, of viewpoint notion is based on CRISP-DM standard through the identification of a set of generic criteria that characterize a viewpoint in KDD. Once instantiated, these criteria define an analyst viewpoint. This viewpoint will guide the execution of the KDD process, and then keep trace of reasoning and major decisions made by the analyst. Then, to formalize interaction and interdependence between various analyses according to different viewpoints, we propose a set of semantic relations between viewpoints based on goal-driven analysis. We have defined equivalence, inclusion, conflict, and requirement relations. These relations allow us to enhance coordination, knowledge sharing and mutual understanding between different actors of a multi-view analysis, and re-usability in terms of viewpoint of successful data mining experiences within an organization. An article selected from the international conference NGNS 2012 [74] will be published in the on-line Journal of Mobile Multimedia, Volume 9 No.3 &4 March 1, 2014.

6.3. Information and Social Networks Mining for Supporting Information Retrieval

6.3.1. Clustering of Relational Data and Social Networks Data: Graph Aggregation

Participant: Yves Lechevallier.

The automatic detection of communities in a social network can provide a kind of graph aggregation. The objective of graph aggregations is to produce small and understandable summaries and it can highlight communities in the network, which greatly facilitates the interpretation.

Social networks allow having a global view of the different actors and different interactions between them, thus facilitating the analysis and information retrieval.

In the enterprise context, a considerable amount of information is stored in relational databases. Therefore, relational database can be a rich source to extract social network.

During this year we updated the program developed by Louati Amine in 2011. A book chapter [34] proposes a new aggregation criteria.

This work is done by Louati Amine (AxIS) in collaboration with Marie-Aude Aufaure, head of the Business Intelligence Team, "Ecole Centrale de Paris", MAS Laboratory.

6.3.2. Multi-View Clustering of Relational Data

Participants: Thierry Despeyroux, Francisco de Carvalho, Yves Lechevallier.

In the work reported in [47] in collaboration with Francisco de A.T. de Carvalho, we introduce an improvement of a clustering algorithm described in [78] that is able to partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. In this version of the prototype clusters depend on the variables of the representation space. These matrices could have been generated using different sets of variables and dissimilarity functions. This method, which is based on the dynamic clustering algorithm for relational data, is designed to provided a partition and a vector of prototypes for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fit between clusters and their representatives. These relevance weights change at each algorithm iteration and are different from one cluster to another. Moreover, various tools for the partition and cluster interpretation furnished by this new algorithm are also presented.

Two experiments demonstrate the usefulness of this clustering method and the merit of the partition and cluster interpretation tools. The first one use a data set from UCI machine learning repository concerning handwritten numbers (digitalized pictures). The second uses a set of reports for which we have an expert classification given a priori. This work has been published this year as a chapter in "Advances in Knowledge Discovery and Management" [32].

6.4. Extension of Usability Methods and Tools

6.4.1. User Evaluation and Tailoring of Personal Information

Participants: Claudia Detraux, Dominique Scapin.

In the context of the ANR project PIMI (Personal Information Management through Internet) an ergonomic evaluation was conducted on the initial prototype, in its PC version [49] and its mobile version [48]. In addition, an experiment was conducted on the usability of the new improved PIMI prototype. The goals were to evaluate its usability, and to assess user tailoring as an evaluation technique. Thirty users participated to the study: a first part consisted in a standard user test (SUT) and a second part was a usability test with tailoring (UTT). Overall, a total of 51 usability problems were diagnosed. Among those, 32 resulted from SUT, and 19 from UTT. Part of the latter (11) are additional to the ones identified during SUT, and to those diagnosed previously by usability inspection (UI with Ergonomic Criteria). The active involvement of users through customization scenarios appear to provide additional cues for usability assessment, and for design, with new generic usability recommendations [23],[22].

6.5. Designing and Evaluating User Experience and Methods for Open Innovation

6.5.1. MyGreenServices: a Pollution Collective-Awareness Platform based on Citizen Sensing

Participants: Brigitte Trousse, Guillaume Pilot, Xavier Augros, Florian Bonacina, Caroline Tiffon, Anne-Laure Negri, Bernard Senach.

Adopting a living lab approach and following an experiential design process [63], we co-created with users and implemented a Pollution Collective-Awareness platform based on Citizen Sensing called "MyGreenServices" [38]. This deployment was very rich in terms of a better understanding of research problems to be addressed in this context in order to lead to user behaviour changes: citizen sensing, environmental crowdsourcing platform and user experience in the context of IoT.

MyGreenServices (http://mygreenservices.inria.fr) which was very robust offers various green services such as the visualization of environmental data collected by citizen, the alert services, the ability to download data, the forum for sharing ideas and best practices in terms of eco-responsible behaviors. MyGreenServices provides access to citizen measures (stations and electric vehicles) for any registered user. Moreover, citizens who host a station can trace the time history of the data sensed. The priority was to provide to users all the IoT data by them. Two ways to represent data have been chosen as shown in Figure 1 :

- The use of maps with measures coming from environmental sensors and based on a colour scale indication;
- The pollution curves that support the cartography and allow the access to the detailed data for the user.

A pollution alert service has been created considering two points of view:

- The first consists of localising a person (with his agreement) and indicating via email or text message the passage through a polluted area;
- The second allow the user to define an area to follow and the user will be advised of pollution alerts for the area by email or text message.

An important effort has been done in designing, testing and improving user interfaces based on pre-test with the usability testing software named Morae and experiments in real situations.

Two experiments have been carried out in February and in June 2013, with the aim to test the platform MyGreenServices by two user profiles (consumers and producers of data) and to measure User experience. The aim of the experiments is to assess the user experience and experiential learning related to MyGreeenServices; this includes experience related to the IoT devices, to the measures and services as well as air quality awareness and behaviour changes monitoring. See Section 6.5.3) for more details on the used model and measurement methodology.

For supporting Citizen Sensing, we elaborated IoT installation guides for our three Pollution stations (based on user feedbacks): Pollux station for dust from CKAB [7], Azimut stations for Ozone and Nitrogen dioxide from Azimut Monitoring [8] and AxISbox stations for dust (Inria Cf. Section 6.5.2).

In order to ensure a proper data analysis, log and usage analytics were structured and gathered in an admin tool designed by the AxIS team at Inria. This tool is a component of the MyGreenServices portal.

6.5.2. AxISbox, a Prototype of a Low-Cost Dust Arduino-based Station

Participant: Guillaume Pilot.

In order to provide more citizen sensors during our Elliot experiments, we developed a first prototype of a new low cost dust (PM10) station (with Rasburry and Arduino) called AxISbox (cf. Figure 2) which we tested for research purposes. This prototype was validated during the second ELLIOT experiment in June.

6.5.3. Modelling and Measuring User Experience for Green IoT-based Services

Participants: Brigitte Trousse, Anne-Laure Negri, Caroline Tiffon, Xavier Augros, Guillaume Pilot.

[7] CKAB URL: http://ckab.com/polluxnz-city
[8] Azimut Monitoring URL: http://www.azimut-monitoring.com/

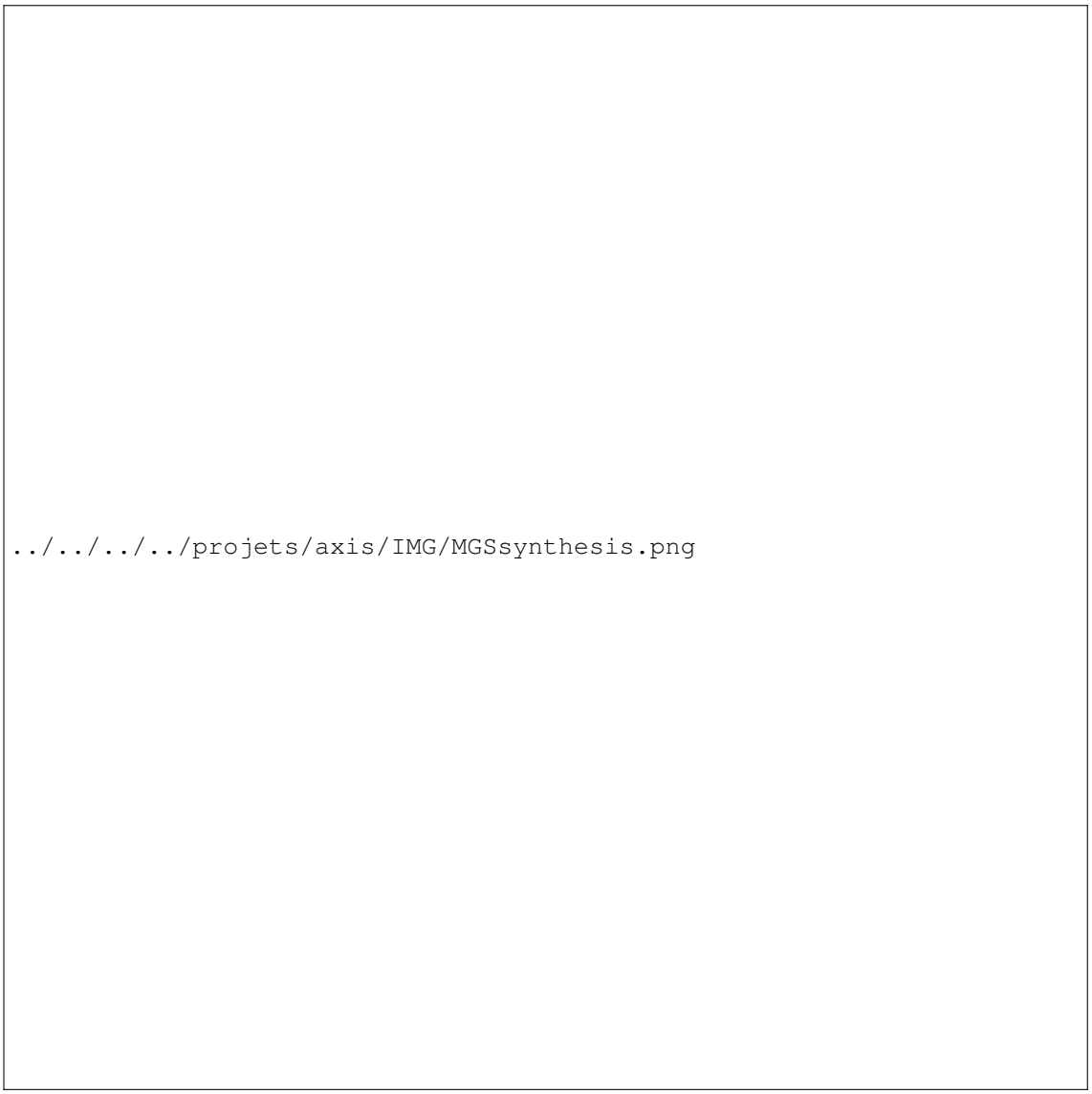


Figure 1. MyGreenServices Platform

Figure 2. Citizen sensors: Pollux station, AxISbox and Azimut mobile station

In accordance with the overall objective of MyGreenServices, we provided an UX modelling and measurement methodology for Green IoT-based services we applied on MyGreenServices. In our ELLIoT context, we focused on the level of awareness/experiential learning raised after usage of MyGreenServices (awareness pollution, awareness of citizen dissemination and change of behaviors), the ease of use and diffusion aspects (as being a tool provided to the citizen). Two objects of the learning were considered: IoT via myGreenServices portal and Air quality. We used a differential between a pre-profile and post-profile. Our UX methodology in the context of ELLIOT project is lying on the five steps we applied on the two versions of MyGreenServices:

- Instantiation of the holistic UX model elaborated within ELLIOT [63] (cf. the first three columns in Figure 3),
- Choice of types of UX momentary, episodic, cumulative) depending on the moment of the measurement (cf. Figure 4),
- Identification of relevant data to be collected and UX indicators (cf. the last two columns in Figure 3),
- Definition of UX metrics for indicators and rules (see Section 6.5.4 for the example of the Usefulness property),
- and finally data pre-processing and UX indicators/properties computation (via for some properties FocusLab 6.6).

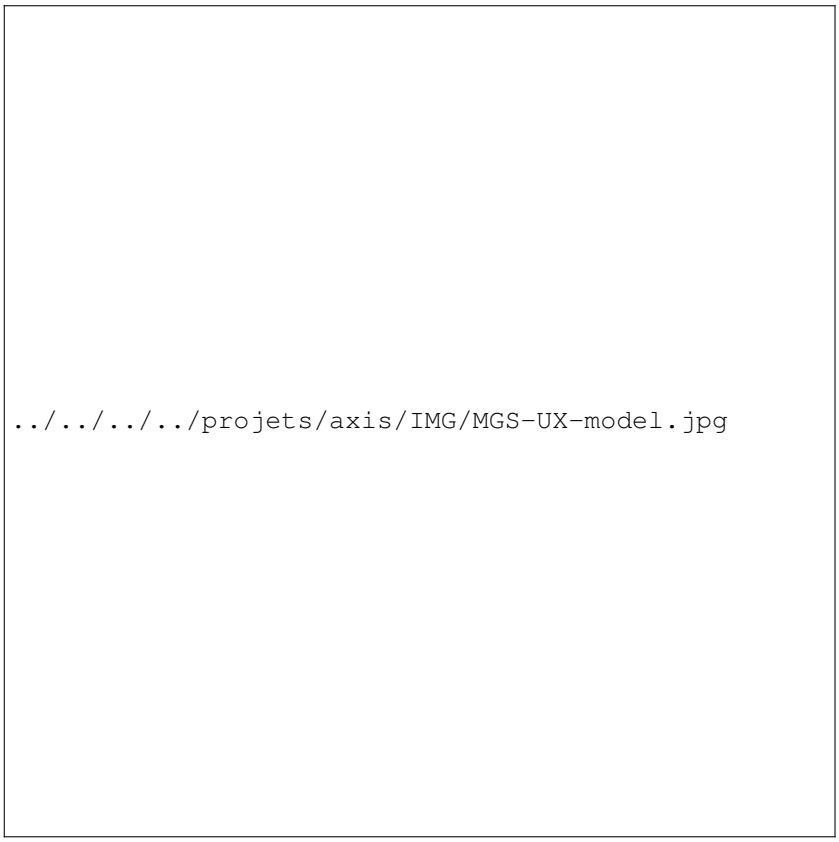


Figure 3. MyGreenServices UX Model

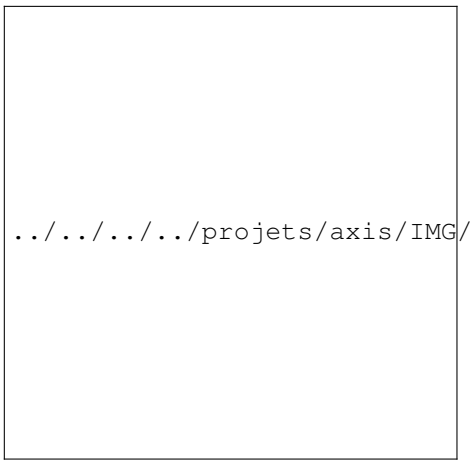


Figure 4. UX Types extracted from [70]

The two experiments clearly indicate both good results in terms of user experience with better result for the second experiment due to the improvement of MyGreenServices (v2) and better community management. A comparative analysis has been made for our two experiments, showing better quantitative value of UX indicators for the second version which was based on User feedback.

6.5.4. Evaluating User Behaviour Changes For MyGreenServices Usefulness Measurement

Participants: Brigitte Trousse, Yves Lechevallier, Xavier Augros, Caroline Tiffon.

The Usefulness UX property of our UX model [38] is calculated by aggregating the analysis of two questions related to a change of behaviours during (4 times) and/or after the experiment in terms of: transportation, aeration, outgoing, sport, aeration or others. We used the web service MNDClustering_Sequence (based on our MND clustering method [45]) to classify the answers to these questions and to provide a sequence of clusters by each user. See Section 6.6.2 related to this new web service.

A data table was built with all the answers for each (user, timestamp) and is analyzed to generate a partition in 3 clusters for the experiment by calling the Focuslab MND webservice (cf. Section 6.2.5) which has been improved this year. The Output via MNDClusterSequence web service is a csv data file with for each user the sequence of 5 clusters obtained during the experiment.

Then we identified the users having changed their behaviour. We use the following UX rules to conclude on this property:

- If % users declaring a change of behaviour $>$ 5% then high
- If % users declaring a change of behaviour $<$ 5% and $>$ 1% then medium
- If % users declaring a change of behaviour $<$ 1% then low

The result is "high" related to our two experiments. Note that other questions related to the usefulness of some MyGreenServices functionalities (alerts, forum, data synthesis, etc.) could be integrated in a more global rule for Usefulness.

6.5.5. Persuasive Technologies in Energy Economy

Participants: Bernard Senach, Anne-Laure Negri.

The ECOFFICES project [51] was for AxIS project team our first step towards eco-behaviour study. This research was complemented in 2012 with a literature review aiming at a deeper understanding of breaks and levers to eco behavior adoption. The work in this topic lead to a presentation [9]in the mobility context during the GreenCode Forum (see the video on YouTube) and to an internal seminar for Axis members. A draft of an Inria research report on this topic has been started.

The two research lines "Energy Economy" and "Persuasive Technology" have been merged and an analysis of the Ecoffices challenge has been engaged in the light of works in the fields of Persuasive Technologies and Game Design. In this analysis, the Ecoffices Energy challenge is considered as an hybrid system combining gamification and persuasive principles. Using available models of each field, the experimental device used in the Ecoffices project is deconstructed and evaluated. The persuasive quality analysis relies on the Persuasive System Design model [62]. Concerning the gaming quality of Ecoffices, a first model (Octalysis http://www.yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/) was discarded and we are now using the gamification principles from the literature for the analysis [76].

At the end of 2012, we joined the work group PISTIL (Persuasive Interaction for SusTainabILity) and engaged several actions within this group and two papers are planned for the JIPS 2014 Special Issues on Persuasive Technologies [10]: one on an analysis of the ECOFFICES challenge (under writing) and another on the design and evaluation of persuasive systems.

6.5.6. *Persuasive Technologies in Green Services*

Participants: Brigitte Trousse, Anne-Laure Negri, Mylène Leitzelman, Florian Bonacina, Caroline Tiffon.

The ELLIOT project was for AxIS project team our second step towards eco-behaviour study. It provided us a very rich context to study behaviour changes related to pollution awareness. Our experimental results showed a very promising tendancy in terms of user behaviour changes and the impact of MyGreenServices on leading user eco-behaviours [38].

Persuasive technologies and gamification were used in the context of green Services use case. A specific focus was on gamification for the two customised Ideastream-based tools we developed for the co-creation step and mainly for the one used inside MyGreenServices platform (see Figure 5).

6.6. FocusLab Platform

6.6.1. *New Graphical Charter and New Functionalities*

Participants: Xavier Augros, Florian Bonacina, Brigitte Trousse.

This year we implemented a new version of the Focuslab platform (v1.3) (http://focuslab.inria.fr) with a new graphical charter, the addition of the documentation part (books, articles, thesis, reports, etc.) and new functionalities such as cross references between the hardware/software parts with the documentation part, the opportunity of reserving hardware, hardware+software or documentation and a new administration interface. This new version has been tested internally in the team at the end of the year.

6.6.2. *FocusLab Generic Web Service: MNDCluster_Sequence*

Participants: Xavier Augros, Yves Lechevallier, Brigitte Trousse.

This year for Elliot purposes, we built a new FocusLab generic Web Service called MNDCluster-Sequence. This web service uses the new release of MND clustering method [80] (cf. Section 6.2.5) which computes the best partition based on all data for each (user, timestamp). Then it builds for each user the sequence of 5 clusters taking into account the five user time stamp in our case. The resulting sequences are then added for each user as new qualified data in the dataset of Green Services.

This web service is added to those already integrated in FocusLab (See for more details our 2012 activity report http://raweb.inria.fr/rapportsactivite/RA2012/axis/uid116.html)

[9]URL: http://www-sop.inria.fr/axis/papers/2012/GreenCode_2012
[10]On-line journal : Journal d' Interaction Personne-Système, Journal of "Association Francophone d'Interaction Homme-Machine".

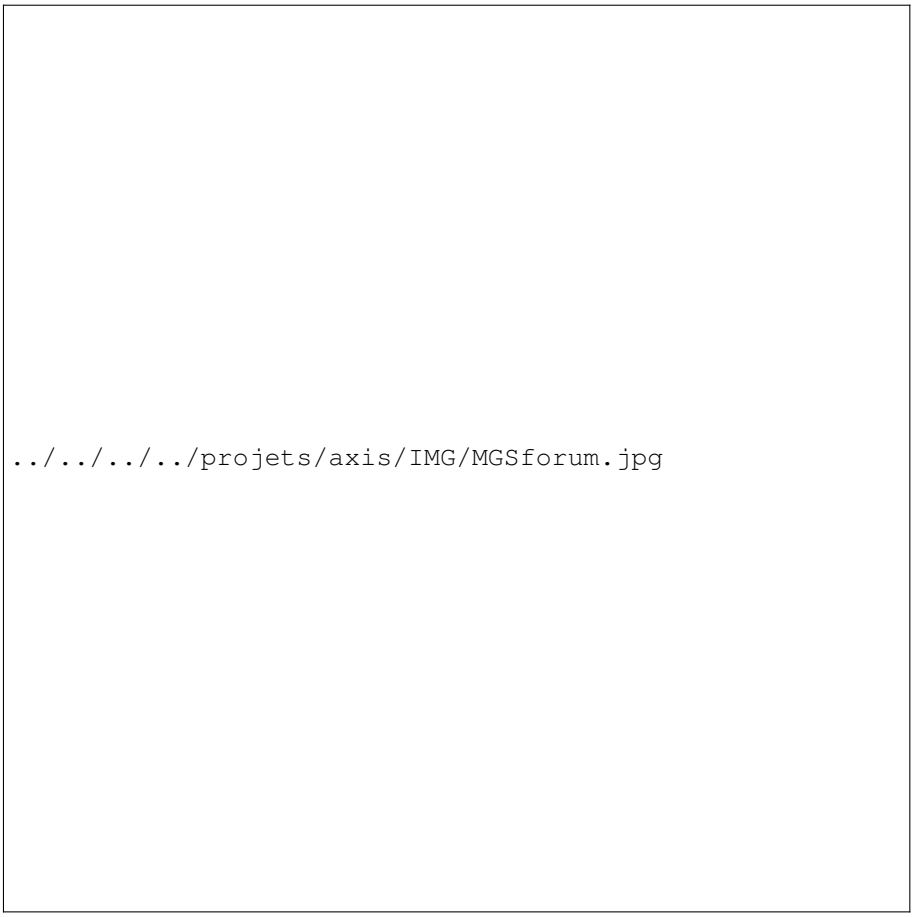


Figure 5. "Gamified Forum" page (including AxISbox)

BANG Project-Team

6. New Results

6.1. Proliferation dynamics and its control

6.1.1. Proliferation dynamics in cell populations

Participants: José Luis Avila Alonso [DISCO project-team, Inria Saclay IdF], Annabelle Ballesta, Gregory Batt [CONTRAINTES project-team], François Bertaux, Frédérique Billy, Frédéric Bonnans [Commands project-team, Inria Saclay IdF], Catherine Bonnet [DISCO project-team, Inria Saclay IdF], Jean Clairambault, Marie Doumic, Xavier Dupuis [Commands project-team], Ján Eliaš, Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Pierre Magal [University Bordeaux II], Anna Marciniak-Czochra [Institute of Applied Mathematics, Universität Heidelberg], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Roberto Natalini [IAC-CNR, Università Sapienza, Rome], Silviu Niculescu [DISCO project-team, Inria Saclay IdF], Hitay Özbay [Bilkent University, Ankara, Turkey], Benoît Perthame, Szymon Stoma [CONTRAINTES project-team], Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Vitaly Volpert [CNRS Lyon, UMR5208, Camille Jordan Institute, Lyon], Jorge Zubelli [IMPA, Rio de Janeiro].

1. **Transition kernels in a McKendrick model of the cell division cycle.** This theme, after a rich harvest of publications (most of them in 2013 and even 2014), is awaiting new developments, since of the main two young researchers on this theme, F. Billy has concluded her 2-year Inria postdoc at Bang, leaving for an industrial company in November 2012, while O. Fercoq (team MaxPlus, Saclay) has defended his PhD thesis at École Polytechnique in September 2012, only to leave for a postdoc position dedicated to optimisation theory in Edinburgh.
2. **Modelling haematopoiesis with applications to AML.** This theme has been active through a collaboration with Inria teams Commands (F. Bonnans, X. Dupuis) and Disco (J.L. Avila Alonso, C. Bonnet, Hitay Özbay, S. Niculescu), and J.-P. Marie's team at St Antoine Hospital leukaemic tumour bank, where A. Ballesta, Cancéropole IdF-Inria postdoc has been detached (ending in January 2013) to identify parameters of a model of acute myeloblastic leukaemia (AML) in patient fresh cell cultures with and without anticancer drugs. This work has led to several presentations, and publications are in preparation. In a book chapter summing up the PhD work of J.L. Avila Alonso [26], and in two submitted conference papers [28], [29], a new model of haematopoiesis for AML is presented, including phases of the cell division cycle and maturation stages, with targets for therapeutic control.
3. **Hybrid models.** Systems combining PDEs and discrete representations in hybrid models, with applications to cancer growth and therapy, in particular for AML, are the object of study of the ANR program *Bimod*, coordinated by V. Volpert (Lyon), associating CNRS (V. Volpert, Lyon), Bordeaux II University (P. Magal) and the Bang project-team.
4. **Molecular model of apoptosis.**

   With G. Gilllet (professor at IBCP/Lyon), A. Ballesta and M. Doumic have designed a mathematical ODE model for the mitochondrial pathway of apoptosis, focused on the early phase of apoptosis (before the cytochrome C release). This model has been validated by experimental data carried out in G. Gillet's lab and applied to propose new therapeutic strategies against cancer [6].
5. **Molecular model of the activity of the p53 protein.** This work, firstly the object of Luna Dimitrio's PhD thesis [37], who left in 2012 for the pharmaceutic industry (SANOFI), has been continued since a new PhD student, Ján Eliaš, has taken over this theme in September 2012 in a new PhD thesis at UPMC, under the supervision of J. Clairambault and B. Perthame. His work has given rise in 2013 to 2 publications [14], [32].

6. **TRAIL - induced apoptosis in HELA cells** Explaining cell-to-cell variability is a major step towards understanding how cancer cells escape action of chemotherapeutic drugs. We set up and studied an integrated model of stochastic gene expression, deterministic translation and protein degradation capable of explaining fractional killing and reversible resistance in Hela cells in response to treatment with TNF-Related Apoptosis Inducing Ligand, TRAIL (Bertaux, Stoma, Drasdo, and Batt, submitted). The results of the model suggests that stochastic fluctuations are a fundamental determinant in understanding cell-to-cell variability, and identified relations between the characteristic time scales of the processes at which stochasticity should play a particular important role.

6.1.2. Physiological and pharmacological control of cell proliferation

Participants: Annabelle Ballesta, Frédérique Billy, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Francis Lévi [INSERM Villejuif (U 776)].

1. *Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type.* This theme (cf. supra "transition kernels...") has been continued [9], [27], [7], [8], [31]. Whereas transition kernels between cell cycle phases without control have been experimentally identified in cell cultures by FUCCI imaging [9], their circadian control remains elusive and has been modelled on the basis of gating by plain cosines representing the influence exerted on these transition kernels by circadian clocks. To go further, it would be necessary to have access by cell imaging to the activity of the best physiological candidates to such gating, namely the cyclin-Cdk complexes, together with the activities of the clock-controlled proteins Wee1 and p21, which thus far have remained unavailable to us through biological experimentation with imaging. A 12-year collaboration work with Francis Lévi on (circadian) chronotherapeutic optimisation in cancer is reported in [30].
2. *Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.* This theme has continued to be developed with new publications for the drugs irinotecan [5], 5-fluorouracil and oxaliplatin [31], and with a recent mini-review by A. Ballesta and J. Clairambault on mathematical models of treatment of metastatic colorectal cancer [4].

6.1.3. Optimisation of cancer chemotherapy and cancer radiotherapy

Participants: Juan Carlos Alfonso [University Complutense, Madrid, Spain], Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Rebecca Chisholm, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Alexandre Escargueil [INSERM and UPMC, St Antoine Hospital], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Miguel Angel Herrero [University Complutense, Madrid, Spain], Michael Hochberg [ISEM, CNRS, Montpellier], Dirk Drasdo, Nick Jagiella, Francis Lévi [INSERM U 776, Villejuif], Thomas Lepoutre [Dracula project-team], Tommaso Lorenzi, Alexander Lorz, Luis Núñez [University Complutense, Madrid, Spain], Benoît Perthame, Emmanuel Trélat [LJLL, UPMC].

1. **Limiting unwanted toxic side effects: age-structured models of the cell cycle.** Optimising cancer chemotherapy, in particular chronotherapy, is the final aim of these activities. A classical numerical method of optimization under the constraint of limiting toxicity to healthy tissues has been applied to the McKendrick model of the cell cycle divided in phases, endowed with physiologically based targets for both internal (circadian) and external (pharmacological) control. This model has been partly biologically identified on continuous FUCCI recordings of proliferating NIH3T3 cells in culture media; these data were made available to us within the C5Sys consortium, an ERASYSBIO+ European project. Then additional theoretical characteristics establishing hypothetical differences between healthy and cancer cell populations, relying on different responses to physiological circadian clock influences on gating by Cyclin-Cdk complexes between cell cycle phases, have been used to solve the optimization problem, proposing an optimal drug infusion regimen [7], [8], [27], [9]. Using an even more complex McKendrick-like model of the cell cycle, a connection with previously established PK-PD ODE models of the anticancer drugs 5-Fluorouracil and Oxaliplatin has

been established, proposing optimized combined drug delivery flows to solve the same optimization problem [31].

2. **Limiting drug resistance in cancer cell populations: cell Darwinism.** This theoretical activity has been continued also in more general settings taking into account another major issue of anticancer treatment, namely resistance to drugs in cancer cells. To this latter aim, we have developed another type of models based on integro-differential equations, which are inspired from those used in ecology for Darwinian evolution [22]. These are aimed at studying another major issue in cancer therapy: appearance of resistances to treatment in tumour cell populations. Indeed, these cell populations, because of their heterogeneity and genomic instability, present an ability to adapt and evolve (in the Darwinian sense) that is much higher than in healthy cell populations [7], [18], [35]. The time scales under investigation, much shorter than in ecology, are however much longer than in microbiology, and are those of clinical treatments. Theoretical optimization of external controls representing combined cytotoxic and cytostatic treatments on these models with the aim to limit the emergence of drug resistance are presently under assessment, in collaboration with Emmanuel Trélat (LJLL, UPMC), paper in preparation.
3. **Molecular aspects: ABC transporters.** From a molecular point of view, studying drug resistance leads to the study of ABC transporters, which is one of the tracks followed by A. Ballesta, following her PhD thesis, in collaboration with F. Lévi's INSERM team in Villejuif [4], [5].
4. **Optimisation of cell kill in AML.** Underway is also the use of methods of optimal control methods developed by the Commands project-team (Frédéric Bonnans, Xavier Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML). X. Dupuis has lately produced a paper [40], accepted for publication in Math. Mod. Phys. Phenom, on optimisation of a combined treatment using a cytotoxic drug (representing Aracytin) and a cytostatic drug (representing AC220, an antagonist of Flt-3 receptors). This work is led in conjunction with the DISCO team, cf. supra "Modelling haematopoiesis with applications to AML").
5. **Estimating dose painting effects in radiotherapy: a mathematical model.** Tumor heterogeneity is widely considered to be a determinant factor in tumor progression and in particular in its recurrence after therapy. Unfortunately, current medical techniques are unable to deduce clinically relevant information about tumor heterogeneity by means of non-invasive methods. As a consequence, when radiotherapy is used as a treatment of choice, radiation dosimetries are prescribed under the assumption that the malignancy targeted is of a homogeneous nature. In this work we discuss the possible effects of different radiation dose distributions on heterogeneous tumors by means of an individual cell-based model. To that end, a case is considered where two tumor cell phenotypes are present, which strongly differ in their respective cell cycle duration and radiosensitivity properties. We show herein that, as a consequence of such differences, the spatial distribution of such phenotypes, as the resulting tumor heterogeneity, can be predicted as growth proceeds. As a consequence, heterogeneous dosimetries can be selected to enhance tumor control by boosting radiation in the region occupied by the more radioresistant tumor cell phenotype. It is also shown that, when compared with homogeneous dose distributions as those being currently delivered in clinical practice, such heterogeneous radiation dosimetries fare always better than their homogeneous counterparts (Alfonso et. al., PLoS One accepted [3]).

6.1.4. Protein polymerisation and application to amyloid diseases

Participants: Annabelle Ballesta, Vincent Calvez [ENS Lyon], Marie Doumic, Pierre Gabriel, Hadjer Wafaâ Haffaf, Benoît Perthame, Stéphanie Prigent [BPCP, INRA Jouy-en-Josas], Human Rezaei [BPCP, INRA Jouy-en-Josas], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

Published in PLoS One in collaboration with the team of biologists led by H. Rezaei [44], a new and very complete PDE model for protein polymerisation has been designed. Following F. Charles's work, A. Ballesta has applied this model to Huntington's disease (PolyQ expansion) and compared it with its ODE counterpart, leading to a better understanding of the leading mechanisms responsible for PolyQ fibrillisation. New applications of this framework model are in progress with H.W. Haffaf and S. Prigent.

The eigenvalue problem playing a major role in the representation of Prion proliferation dynamics and, in a more general way, of many fragmentation-coalescence phenomena, the article [36] investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible nonmonotonic dependency on the parameters, opposite to what would have been conjectured on the basis of some simple cases.

6.1.5. Inverse problem in growth-fragmentation equations

Participants: Marie Doumic, Marc Hoffmann [ENSAE], Nathalie Krell [Univ. Rennes I], Patricia Reynaud [CNRS, Nice Univ.], Lydia Robert [UPMC], Vincent Rivoirard [Paris IX Univ.], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

In collaboration with statisticians (M. Hoffman, Professor at Université de Marne-la-Vallée, V. Rivoirard, MC at Université d'Orsay, and P. Reynaud, CR CNRS at Université de Nice), in the article [38] published in SIAM Num. Anal., we explored a statistical viewpoint on the cell division problem. In contrast to a deterministic inverse problem approach, we take the perspective of statistical inference. By estimating statistically each term of the eigenvalue problem and by suitably inverting a certain linear operator, we are able to construct an estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. Our procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski.

An extension of this work, which consists of the statistical estimation of a branching process modelling the same growth and fragmentation dynamics, has been submitted in [12], in collaboration with N. Krell, M. Hoffmann and L. Robert. Such methods are indeed successfully applied to investigate bacterial growth, in collaboration with L. Robert (INRA and UPMC), see Figure 1 .

In [13], we generalised the inverse techniques proposed previously in [39], [43], in order to adapt them to general fragmentation kernels and growth speeds. The potential applications of this problem are numerous, ranging from polymerisation processes to the cell division cycle. An extension of this work, using refined estimates the Mellin transform of the equation, has just been accepted for publication in *Inverse Problems* [10].

6.2. Tissue growth, regeneration and cell movements

6.2.1. Chemotaxis, self-organisation of cell communities (KPP-Fisher and Keller-Segel)

Participants: Luís Lopes Neves de Almeida, Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], Casimir Emako-Kazianou, François James [univ. Orléans], Alexander Lorz, Grégoire Nadin [UPMC], Benoît Perthame, Jonathan Saragosti [Institut Curie], Pascal Silberzan [Institut Curie], Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet.

Chemotaxis denotes the ability of some cells to undergo a directed movement in response to an extracellular chemical substance. A mathematical description of chemotaxis is a major issue in order to understand collective movements of bacterial colonies. Numerous mathematical models, at various scales, have been proposed, allowing for a good description of cell aggregation under chemotaxis at the macroscopic level, the first of all being that of Keller-Segel (1971), that is now at the centre of an abundant international scientific literature.

At the cell scale, one uses kinetic equations for which numerical simulations have been performed. Behaviour of solutions can be understood by performing a hydrodynamical limit of the kinetic equation. It leads to aggregation type equations for which finite time blow up is observed [42]. Then measure solutions for this system should be considered. A theoretical framework for the existence of weak solutions has then been developed [17], [34] where duality solutions for such system has been investigated which are equivalent to gradient flow solutions [33].

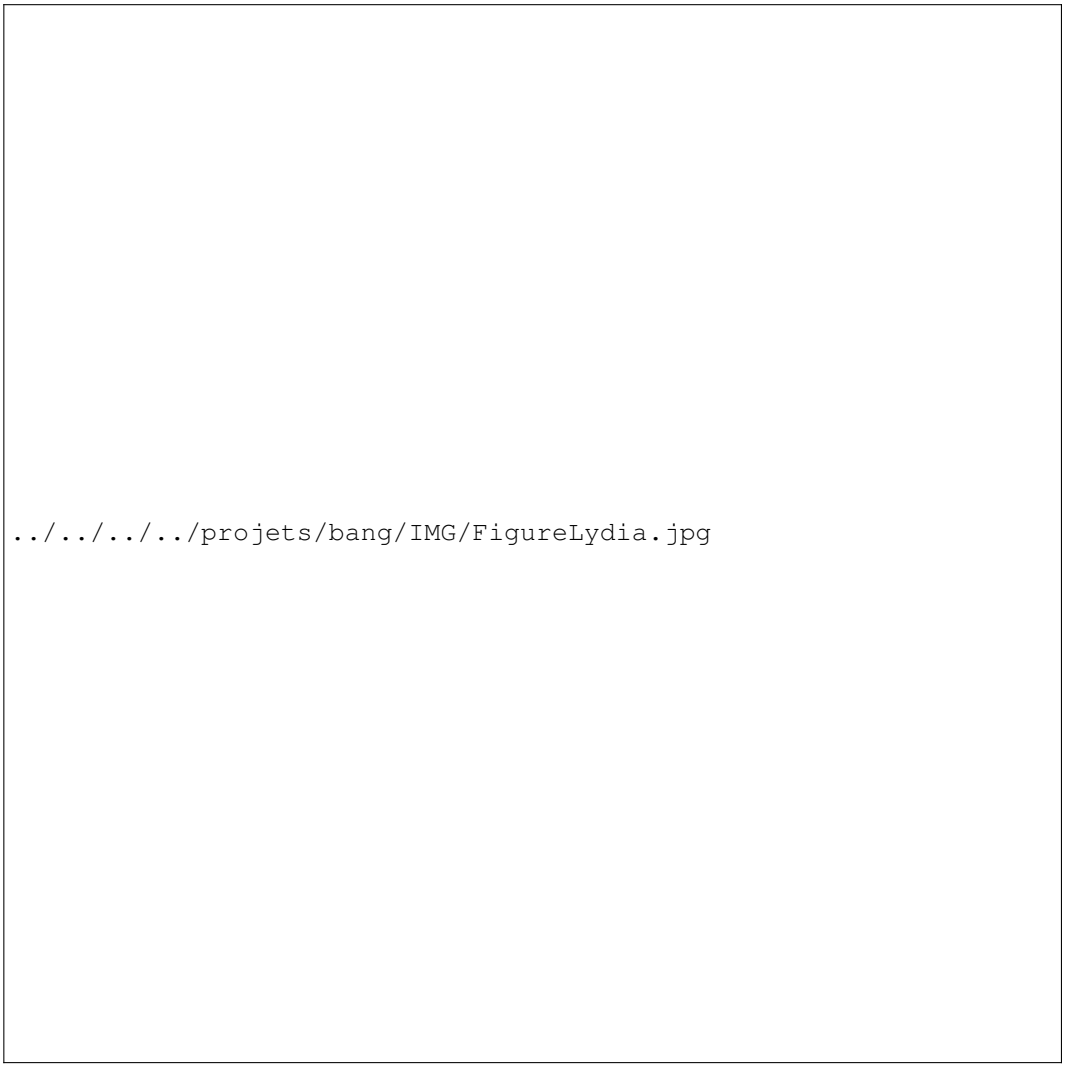


Figure 1. Age and Size Distribution of a bacterial culture (E. coli): comparison between the experimental distribution (A) and the best-fit simulation (B). The methods developed in [38] and [12] allowed us to discriminate between a size-dependent and an age-dependent division rate.

Our understanding of traveling waves has progressed considerably in three directions: fitting continuous models and IBMs [21], fitting precisely models with experiments based on known biological values of parameters, and opening new paradigms: traveling waves can connect a dynamically unstable state to a Turing unstable state, certainly the stable wave connects the unstable state to a pulsating state.

6.2.2. Single-cell-based and continuum models of avascular tumours

Participants: Ibrahim Cheddadi, Dirk Drasdo, Benoît Perthame, Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet, Irène Vignon-Clémentel [REO project-team].

The recent biomechanical theory of cancer growth considers solid tumours as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumour into a host tissue is mainly driven by either diffusion of cells (emerging on the mesoscopic scale by coarse graining from the cell micro-motility) or by cell division depending either on the local cell density (contact inhibition), on mechanical stress in the tumour, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modelling, we prove there are always travelling waves above a minimal speed and we analyse their shapes. They appear to be complex with composite shapes and discontinuities. Several small parameters allow for analytical solutions; in particular the incompressible cells limit is very singular and related to the Hele-Shaw equation. These singular travelling waves are recovered numerically. See [21]. Besides this work, a direct comparison with agent-based and continuum models has been performed, showing very good agreement over a large parameter range.

6.2.3. Single cell-based models of tumour growth, tissue regeneration

Participants: Gregory Batt [CONTRAINTES project-team], François Bertaux, Noémie Boissier, Kai Breuhahn [German Cancer Centre, Heidelberg], Petru Bucur [Hopital Paul Brousse, Paris], Géraldine Cellière, Chadha Chettaoui, Ibrahim Cheddadi, Dirk Drasdo, Adrian Friebel, Rolf Gebhardt [Univ. of Leipzig, Germany], Adriano Henney [Director Virtual Liver Network and VLN consortium], Jan G. Hengstler [Leibniz Research Centre, Dortmund, Germany and CANCERSYS consortium], Stefan Höhme [Research Associate, University of Leipzig], Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Nick Jagiella, Ursula Klingmüller [German Cancer Centre, Heidelberg and LungSys Consortium], Pierre Nassoy [Institut Curie, Paris and Univ. of Bordeaux], Johannes Neitsch, Benoît Perthame, Jens Timmer [University of Leipzig, Germany], Irène Vignon-Clémentel [REO project-team], Paul Van Liedekerke, Eric Vibert [Hôpital Paul Brousse, Villejuif], Ron Weiss [MIT, USA].

1. **Ammonia metabolism in healthy and damaged liver** The model on ammonia detoxification in liver, integrating a compartment model for the glutamine synthetase-active peri-central and the glutamine-inactive peri-portal liver lobule compartment (see Bang report 2012) with the spatial - temporal model of liver regeneration after drug-induced peri-central damage [41] has been extended to include the mass balance of other body compartments. The analysis shows that some body compartments that in the healthy liver produce ammonia, in the damaged liver detoxify blood from ammonia. The detoxification model of liver in combination with the body ammonia balance can be found in ref. (Schliess et. al., Hepatology, accepted [20]).
2. **Drug metabolism in hepatocytes** Since the begining of 2013 animal experiments for testing of cosmetics are forbidden within the EU. This has triggered initiatives towards how modeling may help to investigate drug toxicity, circumventing animal testing. The basic conceptual idea is to test drugs (cosmetics, perspectively also other drugs) in in-vitro systems such as monolayers, sandwich cultures, or multi-cellular spheroids, and use the emerging data to infer the expected toxicity in-vivo using novel experimental and computational approaches [16]. We have integrated an intracellular mathematical model of paracetamol drug metabolism in a mathematical agent-based cell model for monolayer and multi-cellular spheroids and compared simulation results with experimental findings in the same systems. We find that cell-to-cell variabilty can largely explain the experimentally observed cell population survival fractions. The mathematical model is now refined based on measurements of intermediate drug metabolites.

3. **Cell mechanics and its impact on cell proliferation** A novel numerical methodology has been developed to simulate the mechanics of cells and tissues using a continuum approach. Analogously to the Center Based Models, particles are used to represent (parts of) the cells but rather than discrete interactions they represent a continuum. This approach can be used for tissue mechanics simulations in where the individual cell-cell interactions are discarded but instead a constitutive law is proffered [23].
   Moreover, a new model in where cell adhesion dynamics is addressed. The cell model is constructed by a triangulated surface and a coarse-grained internal scaffolding structure. A model cell can adapt to realistic cell shapes, and is able to interact with a substrate or other cells. The parameters in this model can be determined by canonical experiments performed on cells informing about cell deformation, compression and cell-cell adhesion [19].
   A computational model for the confined growth of cells in a capsule has been developed. This model represents a realistic simulation tool for a novel experimental system (Institut Curie, Prof P. Nassoy) in where cells are grown in an elastic environment to mimic the effects of mechanical stress on cells and while monitoring their fate. Model parameter calibration is now ongoing to reproduce the correct quantitative behavior of the cells in order to unravel the relationship between cell mechanical stress and cell behavior.
4. **Playing the game of life with yeast cells** Within a collaboration with a synthetic biology lab at MIT, multicellular modelling of engineered yeast cell populations is performed. Those cells secrete a messenger molecule (IP) which diffuse in the medium, bind to other cells, and trigger a signalling cascade, which finally induces expression of lethal genes. A model has been established based on our single-cell-based model framework associated with PDE simulations, and it is currently used to explain and guide experiments conducted at the MIT. In 2013, the project has achieved significant progress on several aspects. First, we were able to quantitatively reproduce newly produced, rich data on the signaling cascade behavior with a kinetic model describing signaling reactions. Second, comparison between simulations and data allowed to identify key characteristics of the death module, which is positioned downstream of the signaling cascade: there is a rapid and stochastic commitment to death, followed by a deterministic and long delay (2-4 cell generations) needed before cells actually die. Finally, data production and analysis iterations with our collaborators alloweded to optimize the procedures for experimental measurements and the quantitative analysis of data in a synergistic manner.
5. **Other projects in short** Further progress have been achieved on the reconstruction of lung cancer micro-architecture from bright field micrographs. In partial hepatectomy (PHx), pig data on the changes of microarchitecture during regeneration after PHx have been generated and stained now being processed. The image processing chain for liver architecture reconstruction has been refined and extensive analysis has been performed on the architecture of the bile canaliculi network in healthy liver and in disease states of liver. Moreover, non-small-cell lung cancer cell invasion pattern have been analyzed leading to interesting observations now being studied by modelling.
   For multi-scale modeling of liver regeneration after drug-induced pericentral damage, integration of a molecular model of hepatocyte growth factor signalling with an agent-based model of liver regeneration has been extended to include blood flow in the lobule, as well as the contributions of the body compartment to the degradation and production of hepatocyte growth factor (HGF).

6.2.4. Modelling flows in tissues

Participants: Noémie Boissier, Lutz Brusch [TU Dresden], Dirk Drasdo, Adrian Friebel [IZBI, University of Leipzig], Stefan Hoehme [IZBI, University of Leipzig], Nick Jagiella [Inria and IZBI, University of Leipzig], Hans-Ulrich Kauczor [University of Heidelberg, Germany], Fabian Kiessling [University Clinics, Technical University of Aachen, Germany], Ursula Klingmueller [German Cancer Research Centre (DKFZ), Heidelberg, Germany], Hendrik Laue [Fraunhofer Mevis, Bremen, Germany], Ivo Sbazarini [MPI for Molecular Cell Biology and Genetics, Dresden, Germany], Irène Vignon-Clémentel [REO project-team], Marino Zerial [MPI for Molecular Cell Biology and Genetics, Dresden, Germany].

1. **Flow and perfusion scenarios in cancer.** We started reconstruction of the blood vessel system of lung cancers removed by surgery. For this purpose, patients underwent DCE-MRI prior to surgery. Part of the tumors after surgery was sliced and stained for nuclei, proliferation and endothelial cells. The slice data were recorded (Mevis, Luebeck) to allow identification of the position of the indiviudal structures in 3D space. The structures were then segmented. The work turned out to be particularly challenging because of staining artifacts for which image algorithms had to correct for. Nevertheless, last results look promising so that at least the network formed by larger vessels can be segmented and reconstructed in 3D. The so emerging data will be used for modeling of blood flow using the models developed in 2012.
2. **Flow in liver lobules.** We integrated blood flow in the new software CellSys (see above under software) and refined the algorithms. Moreover, we increased the resolution of the capillaries by triangulating them from high resolution confocal scanning micrographs.

6.2.5. Contraction of acto-myosin structures in morphogenesis and tissue repair

Participants: Luís Lopes Neves de Almeida, P. Bagnerini [Univ. Genova], A. Habbal [Univ. Nice], A. Jacinto [CEDOC, Lisbon], M. Novaga [Univ. Padova], A. Chambolle [École Polytechnique].

In 2013 we continued to investigate the dependence of physical and biological mechanisms of actomyosin cable formation and wound closure depending on the geometry of the wound, with particular emphasis on the effect of the wound edge curvature.

When the actomyosin cable starts to contract and the wound starts to close we have noticed that the behavior of the cable is related with the local curvature of the wound edge. This led us to study the curves evolving by positive part of their curvature in a Euclidean framework. A model where we consider viscous behavior and friction in the tissue plus boundary terms associated to cable and lamellipodial forces is under development. The numerical simulations obtained using this model are in good agreement with the previous experimental results and we are pursuing the model development by challenging it with new experiments.

6.3. Neurosciences

Participants: Jonathan Touboul, Gilles Wainrib, Tanguy Cabana, Mathieu Galtier, Luis Garcia Del Molino, Khashayar Pakdaman.

We pursued our studies of disordered networks of the brain and collective phenomena in neuroscience. We have been more interested this year in the role of disorder in the spontaneous emergence of synchronized activity. In order to study these phenomena, we have been establishing limit equations for randomly coupled networks [11], and the analysis of this equation reveal a number of transitions due to the level of disorder in the connectivity. A universal transition observed in such randomly coupled networks is a transition to chaotic activity for large levels of noise. These transitions were investigated [24] and were shown to be related to an explosion of complexity at the edge of chaos, i.e. the number of equilibria is exponentially large with the network size at the phase transition, and the exponential factor was related to the Lyapunov exponent. These large-scale limits give rise to nonlinear reduced equations that we have been introducing in [15]. Eventually, when considering that the network is structured into different populations and that the connectivity weights satisfy a balance condition, which is postulated as a natural scaling of the synaptic input, we have shown that the network shows random transitions to periodic activity depending on the spectrum of the random connectivity matrix [25], yielding up and down states or synchronized oscillations depending on the eigenvalue of larger real part of the connectivity matrix.

CAD Team

5. New Results

5.1. Geometry

5.1.1. From CAD to Engineering: Computing FEM on curved surfaces

Participants: Jean-Claude Paul, Kan-Le Shi, Yu-Shen Liu, Jin-San Cheng, Cheng-Lei Yang, Bruno Durand, Jun-Hai Yong.

In cooperation with Bruno Lévy (Inria)

The cooperation with EADS, based on our new B-Spline surface formulation, was very promising, for complex shape modelling. Our surfaces are very efficient in term of precision. Moreover, they avoid the control point explosion of NURBS surfaces. We propose our work in two directions: 1) to Improve the Modelling process for the user (it is a strategic point of the success of our new mathematical surface); 2) to take profit of the control points way of our surface to compute numerical simulation on this surface directly. In industry, Geometry design and Engineering employ a sequence of tools that are generally not well matched to each other. For example, the output of a computer aided geometric design system is typically not suitable as direct input for a finite-element modeler. This is usually addressed through intermediate tools such as mesh generators. Unfortunately, these are notoriously lacking in robustness. Even once a geometric model has been successfully meshed, the output of a finite-element simulation cannot be directly applied to the original geometric model, since there is no straightforward mapping back to the original design degrees of freedom. Additionally there is a need for a trade-off between the speed of analysis and the fidelity of the results. In the early stages of design, quick results are necessary, but approximate results are acceptable. In the later stages, highly precise results are required, and longer computation times are tolerated. Worse, different underlying models are required for each level of refinement. These difficulties make the design process cumbersome and inhibit rapid iteration over design alternatives. We plan to use FEA on Knot vectors surfaces directly (i.e. use the same function basis for the Geometric Modeling and the Numerical Simulation Process. We will apply this approach to fluids analysis: turbulence modeling (fluid-structure interaction). We think that our surface functions exhibiting higher-order continuity are an ideal candidate for approximating such flows. From the practical point of view, the main objectives of the study are to evaluate, in the scope of this application, the efficiency of such approach in term of simulation accuracy, simulation time and computational convergence. We also aim to evaluation how such approach deals with simulation accuracy/convergence according to CAD definition (quality/size of patches used to define the 3D shape).

5.1.2. From CAD to Manufacturing: Robustness tolerance and error control

Participants: Jun-Hai Yong, Yu-Shen Liu, Clara Issandou, Hai-Chuan Song, Lu Yang, Kang-Lai Qian, Jean-Claude Paul.

In cooperation with Dr. Nabil Anwer – ENS Cachan and the Tsinghua PLM Center (supported by Dassault System). Dr. Yi-Jun Yang (Shandong University), Dr. Xiao-Diao Chen (Zhejiang University)

Based on our theoretical contribution in Differential Geometry, especially about our ϵ-Geometry Continuity and our new geometric operators we proposed several elegant solutions to the most important challenges in Computer Aided Design (see Lees A Piegl. "Ten challenges in Computer-Aided-Design". Jal of CAD 2005. 37 (4): 461-470): robustness, tolerances, error control. During CAD processes one uses a myriad of tolerances, many of which are directly related to the actual manufacturing process. Some interesting questions here include: What are the most relevant machining tolerances? How to set the army of computational tolerances, e.g. those of systems of equations, to guarantee machining within the required accuracy? How tolerances in different spaces, e.g. in model space and in parameter space, are related. Numerical instabilities also account for the majority of computational errors in commercial CAD systems. The problems related to robustness

haunt every programmer who has ever worked on commercial systems. Fixing numerical bugs can be very frustrating, and often times results in patching up the code simply because no solution exists to remedy the problem. We first plan for assisting the designer when specifying the functional tolerances of a single part included in a mechanism, without any required complex function analysis. The mechanism assembly is first described through a positioning table formalism. In order to create datum reference frames and to respect assembly requirements, an ISO based 3D tolerancing scheme will be proposed, thanks to a set of rules based on geometric patterns and TTRS (Technologically and Topologically Related Surfaces). Since it remains impossible to determine tolerance chains automatically, the designer must impose links between the frames. We want to develop proposes ISO based tolerance specifications to help ensure compliance with the designer's intentions, saving on time and eliminating errors.

5.2. Computer Graphics (2010-2013)

5.2.1. Inverse Procedural Modeling of Facade Layouts

Participants: Weiming Dong, Bin Wang, Dong-Ming Yan, Hua-Liang Xie, Jean-Claude Paul.

We want to address the following open research problem: How can we generate a deterministic shape grammar that explains a given facade layout? An approximate dynamic programming framework will tackle this problem. The proposed solution contributes to the compression of urban models, architectural analysis, and the generation of shape grammars for large-scale urban modeling. As a major contribution of this work we want to formulate the inverse procedural modeling problem for facade layouts as a smallest grammar problem. We also want to propose an automatic algorithm to derive a shape grammar for a given facade layout. In this work, we will assume segmented and labeled facade layouts as input and do not derive the shape grammars directly from photographs. The joint optimization of segmentation and grammar extraction remains an aspirational goal for this work.

5.2.2. Architecture Design

Participants: Jean-Claude Paul, Bin Wang, Weiming Dong, Lin Li, Yan Kong, Yong Zhang, Fan Tang, Fuzhang Wu, Cui-Gong Wang.

In cooperation with UC Berkeley - Department of Architecture

We want to propose a method for automated generation of architectural models for computer graphics applications. Our focus is not only on the building layout: the internal organization of spaces within the building, but also the Architectural composition of volumes, roofs and facades. We focus on the generation of various types of buildings: residences, schools, museums, hospitals, civic enters, office buildings. Our work builds on grammar-based procedural modeling, inverse procedural modeling and composition rules, especially symmetry and scaling, and interactivity. Moreover, we consider the architecture design process as an iterative trial-and-error process that requires significant expertise and learning by doing.

CASCADE Project-Team (section vide)

CLASSIC Project-Team

5. New Results

5.1. Contributions earlier to 2013 but only published in 2013

Participants: Gérard Biau, Pierre Gaillard, Gilles Stoltz.

We do not discuss here the contributions provided by [14], [12], [13], [16] since they were achieved in 2012 or earlier (but only published this year due to the reviewing and publishing process).

5.2. Approachability with partial monitoring

Participant: Gilles Stoltz.

This line of research has been developed in our team since its creation (see, in particular, the founding article [9] as well as several other publications in the previous reports). Following the earlier contribution on exhibiting an efficient algorithm for approachability with partial monitoring based on some necessary and sufficient dual condition, we study in [15] the primal approach: the statement of the condition and the existence of (efficient or inefficient) algorithms based on it.

5.3. High-dimensional learning and complex data

Participant: Gérard Biau.

We describe four (not so related) contributions on the theme of high-dimensional learning and complex data.
In [17] we address the problem of supervised classification of Cox process trajectories, whose random intensity is driven by some exogenous random covariable. The classification task is achieved through a regularized convex empirical risk minimization procedure, and a nonasymptotic oracle inequality is derived. The results are obtained by taking advantage of martingale and stochastic calculus arguments, which are natural in this context and fully exploit the functional nature of the problem.
The cellular tree classifier model addresses a fundamental problem in the design of classifiers for a parallel or distributed computing world: Given a data set, is it sufficient to apply a majority rule for classification, or shall one split the data into two or more parts and send each part to a potentially different computer (or cell) for further processing? At first sight, it seems impossible to define with this paradigm a consistent classifier as no cell knows the "original data size", n. However, we show in [18] that this is not so by exhibiting two different consistent classifiers.
A new method for combining several initial estimators of the regression function is introduced. Instead of building a linear or convex optimized combination over a collection of basic estimators $r_1, \cdots, r_M$, [19] uses them as a collective indicator of the proximity between the training data and a test observation. This local distance approach is model-free and very fast. More specifically, the resulting collective estimator is shown to perform asymptotically at least as well in the L^2 sense as the best basic estimator in the collective. A companion R package called COBRA (standing for COmBined Regression Alternative) is presented (downloadable on http://cran.r-project.org/web/packages/COBRA/index.html). Substantial numerical evidence is provided on both synthetic and real data sets to assess the excellent performance and velocity of the method in a large variety of prediction problems.
The impact of letting the dimension d go to infinity on the L^p-norm of a random vector with i.i.d. components has surprising consequences, which may dramatically affect high-dimensional data processing. This effect is usually referred to as the *distance concentration phenomenon* in the computational learning literature. Despite a growing interest in this important question, previous work has essentially characterized the problem in terms of numerical experiments and incomplete mathematical statements. In the paper [20], we solidify some of the arguments which previously appeared in the literature and offer new insights into the phenomenon.

5.4. Dimension free principal component analysis

Participants: Olivier Catoni, Ilaria Giulini.

In a work in progress, Ilaria Giulini, as part of her PhD studies, proved the following dimension free inequality, related to Principal Component Analysis in high dimension. Given an i.i.d. sample X_i, $1 \leq i \leq n$ of vector valued random variables $X_i \in \mathbf{R}^d$, there exists an estimator $\widehat{N}$ of the quadratic form $N(\theta) = \mathbf{E}(\langle \theta, X \rangle^2)$ such that for any $n \leq 10^{20}$, with probability at least $1 - 2\epsilon$, for any $\theta \in \mathbf{R}^d$,

$$\mathbf{1}\left(4\mu < 1\right) \left| \frac{\widehat{N}(\theta)}{N(\theta)} - 1 \right| \leq \frac{\mu}{1 - 4\mu},$$

where

$$\mu = \sqrt{\frac{2.07\ (\kappa - 1)}{n} \left[\log\left(\epsilon^{-1}\right) + 4.3 + \frac{1.6\ \|\theta\|^2 \mathbf{Tr}(G)}{N(\theta)} \right]} + \sqrt{\frac{184\ \kappa \|\theta\|^2 \mathbf{Tr}(G)}{nN(\theta)}},$$

where $G = \mathbf{E}\left(XX^\top\right)$ is the Gram matrix and where $\kappa = \sup \left\{ \dfrac{\mathbf{E}\left(\langle \theta, X \rangle^4\right)}{\mathbf{E}\left(\langle \theta, X \rangle^2\right)^2}, \theta \in \mathbf{R}^d \smallsetminus \mathbf{Ker}(G) \right\}$ is some kurtosis coefficient. This result proves that the expected energy in direction θ can be estimated at a rate that is independent of the dimension of the ambient space $\mathbf{R}^d$. It is obtained using PAC-Bayes inequalities with Gaussian parameter perturbations. The same bound holds in a Hilbert space of infinite dimension, opening the possibility of a rigorous mathematical study of kernel principal component analysis of random data, where the data are represented in a possibly infinite dimensional reproducing kernel Hilbert space.

5.5. Statistical models for corpus linguistics

Participants: Olivier Catoni, Thomas Mainguy.

In [21] we describe a language model as the invariant measure of a Markov chain on sentence samples. The kernel of this Markov chain is defined with the help of some context free grammars : from the sentence sample, a random parse model produces a context free grammar with weighted rules, and from this grammar, a new sentence sample is formed by applying the rules randomly. We prove various mathematical properties of this Markov process, related to its computation cost and the fact that it is weakly reversible and therefore ergodic on each of its communicating classes. As a companion to the Markov chain on sentence samples, we can also define a Markov chain on weighted context free grammars. This leads to another type of grammar, that we called Toric Grammars, defined by a family of context tree grammars that can be computed from any of its members as the communicating class of a Markov chain on context free grammars with weighted rules. Preliminary simulations on small data sets are very encouraging, in that they show that this type of model is able to grasp the recursive nature of natural languages.

CLIME Project-Team

6. New Results

6.1. New methods for data assimilation

One major objective of Clime is the conception of new techniques for data assimilation in geophysical sciences. Clime is active on several of the most challenging theoretical aspects of data assimilation: data assimilation methods based on non-Gaussian assumptions, methods for estimating errors, ensemble filtering techniques, 4D variational assimilation approaches, ensemble-variational methods, etc.

This year, focus was on ensemble-variational methods. We introduced a new method known as the iterative ensemble Kalman smoother. It is an ensemble method with an underlying cost function; it does not require the use of the adjoint; and it is flow-dependent. Because of these propreties, the IEnKS outperforms other data assimilation methods when tested with perfect meteorological toy-models. Its potential for parameter estimation has also been demonstrated.

6.1.1. An iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

The iterative ensemble Kalman filter (IEnKF) was recently proposed to improve the performance of ensemble Kalman filtering with strongly nonlinear geophysical models. IEnKF can be used as a lag-one smoother and extended to a fixed-lag smoother: the iterative ensemble Kalman smoother (IEnKS [12]). IEnKS is an ensemble variational method. It does not require the use of the tangent linear of the evolution and observation models, nor the adjoint of these models: the required sensitivities (gradient and Hessian) are obtained from the ensemble. Looking for the optimal performance, we consider a quasi-static algorithm, out of the many possible extensions. IEnKS is explored on the Lorenz'95 model and on a 2D turbulence model. As a logical extension of IEnKF, IEnKS significantly outperforms standard Kalman filters and smoothers in strongly nonlinear regimes. In mildly nonlinear regimes (typically synoptic scale meteorology), its filtering performance is marginally but clearly better than the standard ensemble Kalman filter, and it keeps improving as the length of the temporal data assimilation window is increased. For long windows, its smoothing performance very significantly outranks the standard smoothers, which is believed to stem from the variational but flow-dependent nature of the algorithm. For very long windows, the use of a multiple data assimilation variant of the scheme, where observations are assimilated several times, is advocated. This paves the way for finer re-analysis freed from the static prior assumption of 4D-Var, but also partially freed from the Gaussian assumptions that usually impede standard ensemble Kalman filtering and smoothing.

6.1.2. Joint state and parameter estimation with an iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

Both ensemble filtering and variational data assimilation methods have proven being useful in the joint estimation of state variables and parameters of geophysical models. Yet, their respective benefits and drawbacks in this task are distinct. An ensemble variational method, known as the iterative ensemble Kalman smoother (IEnKS), has recently been introduced. It is based on an adjoint-free variational but flow-dependent scheme. As such, IEnKS is a candidate tool for joint state and parameter estimation that may inherit the benefits from both the ensemble filtering and variational approaches.

In this study [13], an augmented state IEnKS is tested on the estimation of the forcing parameter of the Lorenz'95 model. Since joint state and parameter estimation is especially useful in applications where the forcings are uncertain but nevertheless determining, typically in atmospheric chemistry, the augmented state IEnKS is tested on a new low-order model that combines the Lorenz'95 model, representing its meteorological part, and the advection diffusion of a tracer for its chemical part. In these experiments, IEnKS is compared to the ensemble Kalman filter, the ensemble Kalman smoother and a 4D-Var method, that are considered choices to solve these joint estimation problems. In this low-order model context, IEnKS is shown to significantly outperform those methods, for any length of the data assimilation window, and for present time analysis as well as retrospective analysis. Besides, the performance of IEnKS is even more striking on parameter estimation, whereas getting close to the same performance with 4D-Var is likely to require both a long data assimilation window and a complex modeling of the background statistics.

6.1.3. Data assimilation applied to air quality at urban scale

Participants: Vivien Mallet, Raphaël Périllat, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Frédéric Mahé [Airparif], Pierre Pernot [Airparif], Fabrice Joly [Airparif].

Based on Verdandi [14], Polyphemus and the "Urban Air Quality Analysis" software, data assimilation was further developed at urban scale. The Best Linear Unbiased Estimator (BLUE) is computed to merge the outputs of the ADMS Urban model and the observations of a sparse monitoring network [19]. We improved the modeling of the covariance of the model state error. The assimilation was applied for part of Paris (see Fig. 2) and for Paris region, in the context of the PRIMEQUAL project PREQUALIF ("Multidisciplinary Program on Air Quality Research in Île-de-France").

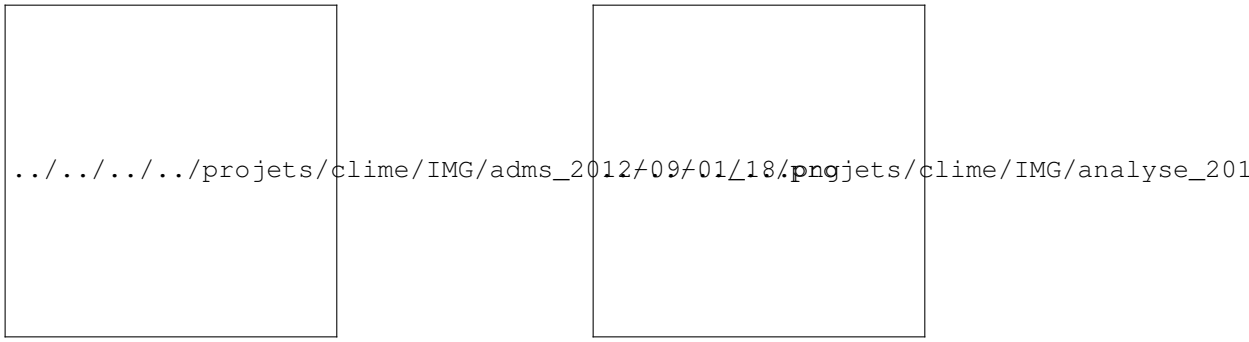


Figure 2. Left: Map of [NO_2] ($\mu g\ m^{-3}$), before assimilation, at a given date in September 2012. Right: Map of [NO_2] ($\mu g\ m^{-3}$), after assimilation of the observations (disks).

It was applied to nitrogen dioxide, particulate matter and black carbon. Specific investigations were carried out to estimate the variance of the a posteriori error and to determine the impact of each monitoring station on the final results.

6.2. Inverse modeling

We continued research on inverse modelling techniques, with a focus on hyperparameter estimation when the statistics are non-Gaussian. We applied these methods to the estimation of the caesium-137 Fukushima source term using heterogenous datasets. We applied similar methods to the estimation of Volatile Organic Compounds (VOC) at the European scale by assimilation of the EMEP VOC observations over one year. We also studied the estimation of several hyperparameters in the context of CO_2 flux inversions.

6.2.1. Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations

Participants: Victor Winiarek, Marc Bocquet, Nora Duhanyan [CEREA], Yelva Roustan [CEREA], Olivier Saunier [IRSN], Anne Mathieu [IRSN].

To estimate the amount of radionuclides and the temporal profile of the source term released in the atmosphere during the accident of the Fukushima Daiichi nuclear power plant in March 2011, inverse modeling techniques have been used and have proven their ability in this context. In a previous study, the lower bounds of the caesium-137 and iodine-131 source terms were estimated with such techniques, using activity concentration observations. The importance of an objective assessment of prior errors (the observation errors and the background errors) was emphasised for a reliable inversion. In such critical context where the meteorological conditions can make the source term partly unobservable and where only a few observations are available, such prior estimation techniques are mandatory, the retrieved source term being very sensitive to this estimation.

We propose to extend the use of these techniques to the estimation of prior errors when assimilating observations from several data sets [21]. The aim is to compute an estimate of the caesium-137 source term jointly using all available data about this radionuclide, such as activity concentrations in the air, but also daily fallout measurements and total cumulated fallout measurements. It is crucial to properly and simultaneously estimate the background errors and the prior errors relative to each data set. A proper estimation of prior errors is also a necessary condition to reliably estimate the a posteriori uncertainty of the estimated source term. Using such techniques, we retrieve a total released quantity of caesium-137 in the interval $11.6 - 19.3$ PBq with an estimated standard deviation range of $15 - 20\%$ depending on the method and the data sets. The "blind" time intervals of the source term have also been strongly mitigated compared to the first estimations with only activity concentration data.

6.2.2. An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations

Participants: Olivier Saunier [IRSN], Anne Mathieu [IRSN], Damien Didier [IRSN], Maryline Tombette [IRSN], Denis Quélo [IRSN], Victor Winiarek, Marc Bocquet.

The Chernobyl nuclear accident, and more recently the Fukushima accident, highlighted that the largest source of error on consequences assessment is the source term, including the time evolution of the release rate and its distribution between radioisotopes. Inverse modeling methods, which combine environmental measurements and atmospheric dispersion models, have proven being efficient in assessing source term due to an accidental situation. Most existing approaches are designed to use air sampling measurements and some of them also use deposition measurements [21]. Some studies have been conceived to use dose rate measurements, but none of the developed methods were carried out to assess the complex source term of a real accident situation like the Fukushima accident. However, dose rate measurements are generated by the most widespread measurement system and, in the event of a nuclear accident, these data constitute the main source of measurements of the plume and radioactive fallout during releases. This study [18], [23] proposes a method to use dose rate measurements as part of an inverse modeling approach to assess source terms. The method is proven efficient and reliable when applied to the accident at the Fukushima Daiichi Nuclear Power Plant (FD-NPP). The emissions for the eight main isotopes have been assessed. Accordingly, 105.9 PBq of ^{131}I, 35.8 PBq of ^{132}I, 15.5 PBq of ^{137}Cs and 12,134 PBq of noble gases were released. The events at FD-NPP (such as venting, explosions, etc.) known to have caused atmospheric releases are well identified in the retrieved source term. The estimated source term is validated by comparing simulations of atmospheric dispersion and deposition with environmental observations. In total, it was found that for 80 % of the measurements, simulated and observed dose rates agreed within a factor of 2. Changes in dose rates over time have been overall properly reconstructed, especially in the most contaminated areas to the northwest and south of the FD-NPP. A comparison with observed atmospheric activity concentration and surface deposition shows that the emissions of caesiums and ^{131}I are realistic but that ^{132}I and ^{132}Te are probably underestimated and noble gases are likely overestimated. Finally, an important outcome of this study is that the method proved to be

perfectly suited to emergency management and could contribute to improve emergency response in the event of a nuclear accident.

6.2.3. Estimation of volatile organic compound emissions for Europe using data assimilation

Participants: Mohammad Reza Koohkan, Marc Bocquet, Yelva Roustan [CEREA], Yougseob Kim [CEREA], Christian Seigneur [CEREA].

The emissions of non-methane volatile organic compounds (VOCs) over western Europe for the year 2005 are estimated via inverse modeling by assimilation of in situ observations of concentration and they are subsequently compared to a standard emission inventory. The study [16] focuses on fifteen VOC species: five aromatics, six alkanes, two alkenes, one alkyne and one biogenic diene. The inversion relies on a validated fast adjoint of the chemical transport model used to simulate the fate and transport of these VOCs. The assimilated ground-based measurements over Europe are provided by the European Monitoring and Evaluation Programme (EMEP) network. The background emissions errors and the prior observational errors are estimated by maximum likelihood approaches. The positivity assumption on the VOC emission fluxes is pivotal for a successful inversion and this maximum likelihood approach consistently accounts for the positivity of the fluxes. For most species, the retrieved emissions lead to a significant reduction of the bias, which underlines the misfit between the standard inventories and the observed concentrations. The results are validated through a forecast test and a cross-validation test. An estimation of the posterior uncertainty is also provided. It is shown that the statistically consistent non-Gaussian approach, based on a reliable estimation of the errors, offers the best performance. The efficiency in correcting the inventory depends on the lifetime of the VOCs and the accuracy of the boundary conditions. In particular, it is shown that the use of in situ observations using a sparse monitoring network to estimate emissions of isoprene is inadequate because its short chemical lifetime significantly limits the spatial radius of influence of the monitoring data. For species with longer lifetime (a few days), successful, albeit partial, emission corrections can reach regions hundreds of kilometres away from the stations. Domainwide corrections of the emissions inventories of some VOCs are significant, with underestimations on the order of a factor of two for propane, ethane, ethylene and acetylene.

6.2.4. Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions

Participants: Lin Wu [LSCE, France], Marc Bocquet, Frédéric Chevallier [LSCE, France], Thomas Lauvaux [Department of Meteorology, Pennsylvania State University, USA], Kenneth Davies [Department of Meteorology, Pennsylvania State University, USA].

Uncertainty quantification is critical in the inversion of CO_2 surface fluxes from atmospheric concentration measurements. We estimate the main hyperparameters of the error covariance matrices for a priori fluxes and CO_2 concentrations, that is, the variances and the correlation lengths, using real, continuous hourly CO_2 concentration data in the context of the Ring 2 experiment of the North American Carbon Program Mid Continent Intensive. Several criteria, namely maximum likelihood (ML), general cross-validation (GCV) and χ^2 test are compared for the first time under a realistic setting in a mesoscale CO_2 inversion. It is shown [22] that the optimal hyperparameters under the ML criterion assure perfect χ^2 consistency of the inverted fluxes. Inversions using the ML error variances estimates rather than the prescribed default values are less weighted by the observations, because the default values underestimate the model-data mismatch error, which is assumed to be dominated by the atmospheric transport error. As for the spatial correlation length in prior flux errors, the Ring 2 network is sparse for GCV and this method fails to reach an optimum. In contrast, the ML estimate (e.g. an optimum of 20 km for the first week of June 2007) does not support long spatial correlations that are usually assumed in the default values.

6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring networks and to new monitoring strategies. This year, we studied the impact of using lidar observation for particulate matter forecasting.

6.3.1. Assimilation of ground versus lidar observations for PM_{10} forecasting

Participants: Yiguo Wang [CEREA], Karine Sartelet [CEREA], Marc Bocquet, Patrick Chazette [LSCE, France].

This study [20] investigates the potential impact of future ground-based lidar networks on analysis and short-term forecasts of PM_{10}. To do so, an Observing System Simulation Experiment (OSSE) is built for PM_{10} data assimilation using optimal interpolation over Europe for one month in 2001. First, we estimate the efficiency of the assimilation of lidar network measurements in improving PM_{10} concentration analysis and forecast. It is compared to the efficiency of assimilating concentration measurements from the AirBase ground network, which includes about 500 stations in western Europe. It is found that the assimilation of lidar observations is more efficient at improving PM_{10} concentrations in terms of root mean square error and correlation after 12 hours of assimilation than the assimilation of AirBase measurements. Moreover, the spatial and temporal influence of the assimilation of lidar observations is larger and longer. In our experiments, the assimilation of lidar products improves PM_{10} forecast for 108 hours against 60 hours for AirBase assimilation. The results show a potentially powerful impact of the future lidar networks. Secondly, since a lidar is a very costly instrument, a sensitivity study on the number of required lidars is performed to help defining an optimal lidar network for PM_{10} forecast. The results suggest 12 lidar stations over western Europe, because a network with 26 lidar stations is more expensive and offers a limited improvement (less than 1 $\mu g\ m^{-3}$ of root mean square error on average) over the lidar network. A comparison of two networks with 12 lidar stations at different locations does not lead to substantial differences.

6.4. Reduction and emulation

The use of environmental models raise 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.

In particular, the application of data assimilation methods and uncertainty quantification techniques may require dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction can be carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.4.1. Reduction and emulation of a chemistry-transport model

Participants: Vivien Mallet, Serge Guillas [University College London].

Both reduction and emulation were applied to the dynamic air quality model Polair3D from Polyphemus. The reduction relied on proper orthogonal decomposition (POD) on the input data and on the state vector. The dimension of the reduced subspace for the input data is about 80, while the dimension of the reduced state vector is less than 10. The projection of the state vector on its reduced subspace can be carried out before every integration time step, so that one can reproduce a full state trajectory (in time) using the reduced model.

Significant advances were made to emulate the reduced model, which requires about 90 inputs (reduced input data and reduced state vector) and computes about 10 outputs (reduced state vector). 90 inputs is however a large number to build an emulator using a classical approaches. Promising results were however obtained with radial basis functions and an adapted kriging-based method.

6.4.2. Reduction and emulation of a static air quality model

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech].

The dimension reduction was applied to the outputs of the urban air quality model ADMS Urban, which is a static model with low-dimensional inputs and high-dimensional outputs. A proper orthogonal decomposition (POD) on the outputs allowed us to drastically reduce their dimension, from 10^4 to just a few scalars. The emulation of the reduced model itself was successfully carried out with radial basis functions or an adapted kriging-based method. The resulting reduced/emulated model exhibited meaningful response to all variables. Its performance compared to observations was the same as the original model. The computational cost of the full model is about 8 minutes on 16 cores (for a single time step), while the reduced/emulated model requires only 50 ms on one core [29].

6.4.3. Motion estimation from images with a waveforms reduced model

Participants: Etienne Huot, Isabelle Herlin, Giuseppe Papari [Lithicon, Norway], Karim Drifi.

Dimension reduction is applied to an image model, composed of Lagrangian constancy of velocity and transport of image brightness. Waveforms basis are obtained on the image domain for subspaces of images, motion fields and divergence-free motion fields, as eigenvectors of quadratic functions. Image assimilation with th reduced model allows to estimate velocity fields satisfying space-time properties defined by user and traduced as a quadratic function. This approach also solves the issue of complex geographical domains and the difficulty of applying boundary conditions on these domains. Results are obtained with a reduced dimension of motion to a few scalars, to be compared with the original problem that has the size of image domain [31], [26], [25].

6.5. Ensemble forecasting with sequential aggregation

The aggregation of an ensemble of forecasts is an approach where the members of an ensemble are given a weight before every forecast time, and where the corresponding weighted linear combination of the forecasts provides an improved forecast. A robust aggregation can be carried out so as to guarantee that the aggregated forecast performs better, in the long run, than any linear combination of the ensemble members with time-independent weights. The approaches are then based on machine learning. The aggregation algorithms can be applied to forecast analyses (generated from a data assimilation system), so that the aggregated forecasts are naturally multivariate fields.

6.5.1. Application of sequential aggregation to meteorology

Participants: Paul Baudin, Vivien Mallet, Gilles Stoltz [CNRS], Laurent Descamps [Météo France].

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. Ensembles of forecasts for wind velocity and mean sea level pressure, from Météo France, were aggregated. Preliminary results show significant improvements for mean sea level pressure.

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Sergiy Zhuk [IBM research, Ireland], Paul Baudin, Gilles Stoltz [CNRS].

An important issue is the estimation of the uncertainties associated with the aggregated forecasts. One investigated direction relies on the framework of machine learning, with the aggregation of an ensemble of probability density functions instead of the point forecasts of the ensemble.

Another direction is to reformulate the aggregation problem in a filtering problem for the weights. The weights are supposed to satisfy some dynamics with unknown model error, which defines the state equation of a filter. An observation equation compares the aggregated forecast with the observations (or analyses) with known observational error variance. The filter finally computes estimates for the weights and quantifies their uncertainties. We applied a Kalman filter and a minimax filter for air quality forecasting. We also introduced a criterion that the filter results should satisfy if they are representative of the uncertainties [17].

6.6. Uncertainty quantification

Many uncertainties limit the forecast skills of geophysical simulations: limited understanding of physical phenomena, simplified representation of a system state and of the physical processes, inaccurate data and approximate numerical solutions. In many applications, a deterministic forecast or analysis is not enough a result since its uncertainties may be very large. It is of high interest to evaluate the quality of a forecast, before observations are available, and the quality of an analysis at any location, observed or not. An even more desirable result is the full probability density of system state, which can only be derived from a fully stochastic approach.

6.6.1. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard [IRSN], Vivien Mallet, Irène Korsakissok [IRSN].

We carried 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. The sensitivities to most input parameters were computed using the Morris method (with 8 levels and 100 trajectories). The influences of 19 scalar parameters were quantified. The scalar parameters were additive terms or multiplicative factors applied to 1D, 2D or 3D fields such as emission rates, precipitations, cloud height, wind velocity. It was shown that, depending on the output quantities of interest (various aggregated atmospheric and ground dose rates), the sensitivity to the inputs may greatly vary. Very few parameters show low sensitivity in any case. The vertical diffusion coefficient, the scavenging factors, the winds and precipitation intensity were found to be the most influential inputs. Most input variables related to the source term (emission rates, emission dates) also had a strong influence.

6.7. 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, these 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.

6.7.1. Divergence-free motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin, Sergiy Zhuk [IBM Research, Ireland].

This research addresses the issue of divergence-free motion estimation on an image sequence, acquired over a given temporal window. Unlike most state-of-the-art technics, which constrain the divergence to be small thanks to Tikhonov regularization terms, a method that imposes a null value of divergence of the estimated motion is defined.

Motion is either characterized by its vorticity value or by its coefficients on a divergence-free basis and assumed to satisfy the Lagragian constancy hypothesis. An image assimilation method, based on the 4D-Var technic, is defined that estimates motion as a compromise between the evolution equations of vorticity or projection coefficients and the observed sequence of images.

The method is applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface.

6.7.2. Model error and motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

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 simultaneously depends on the initial motion field, at the begining 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 has been used to estimate the impact of geophysical forces (gravity, Coriolis, diffusion) and better assess the surface dynamics [24].

6.7.3. Tracking of structures from an image sequence

Participants: Yann Lepoittevin, Isabelle Herlin, 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, and those of the distance map modeling the tracked structures. The method is applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion, as seen on Fig. 3 .

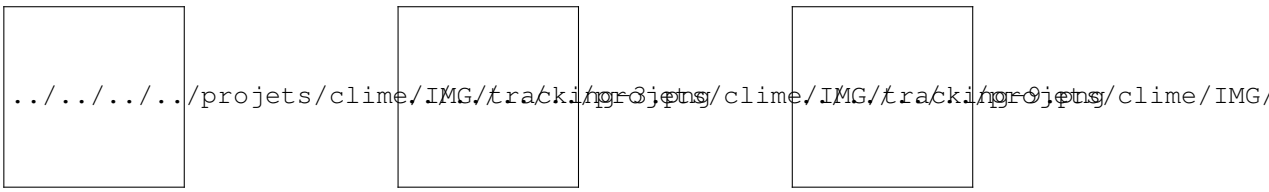


Figure 3. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

Quantification is obtained on synthetic experiments by comparing trajectories of characteristic points. The respective position of these points on the last image of the sequence for different methods may be compared to that obtained with ground truth as seen on Fig. 4 .

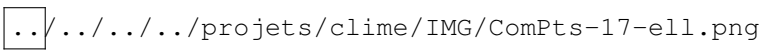


Figure 4. Red point: ground truth. Blue point: our method. Green point: Sun's optical flow. Blue ellipse: our method is the best. Green ellipse: Sun's result is the best. Grey ellipse : results are equivalent.

Data assimilation is performed either with a 4D-Var variational approach [27], [30], [28] or with an ensemble approach. In the last case, computation of the ensemble from optical flow methods of the literature is currently studied.

6.8. Minimax filtering

In minimax filtering for state estimation, the initial state error, the model error and the observation errors are supposed to belong to one joint ellipsoid. It is only assumed that the errors, stochastic or deterministic, are bounded. During the assimilation process, the filter computes an ellipsoid where one will find at least all

states compatible with observations and errors description. The state estimate is taken as the center of the ellipsoid. No assumption on the actual distribution of the errors in needed and the state estimate minimizes the worst-case error, which makes the filter robust.

6.8.1. Retrieval of a continuous image function and a posteriori minimax motion estimation

Participants: Sergiy Zhuk [IBM Research, Ireland], Isabelle Herlin, Olexander Nakonechnyi [Taras Shevchenko National University of Kyiv], Jason Frank [CWI, the Netherlands].

An iterative minimax method is developed for the problem of motion estimation from an image sequence. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the velocity. The algorithm consists of the following parts:

1) we construct a continuous image function $\widehat{I}$, solving the optical flow constraint, such that $\widehat{I}$ fits (in the sense of least-squares) the observed sequence of images. To do so, we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field $\mathbf{w}$, obtained at the previous iteration of the algorithm, and construct the minimax estimate $\widehat{I}$ of the resulting linear advection equation using the observed image sequence as discrete measurements of the brightness function;

2) we plug the estimate of the image gradient, obtained out of pseudo-observations $\widehat{I}$ in 1), into the optical flow constraint and the current minimax estimate $\mathbf{w}$ of the velocity field into the non linear part of Navier-Stokes equations so that we end up with a system of linear PDEs, which represents an extended state equation: it contains a linear parabolic equation for the velocity field and linear advection equation for the image brightness function. We construct the minimax estimate of the velocity field from the extended state equation using again the observed image sequence as discrete measurements of the brightness function;

3) we use the minimax estimate of the velocity field obtained in 2) in order to start 1) again.

Alternatively, point 1) may be used to retrieve a continuous image function from sparse and noisy image snapshots, based on previous motion estimation with a 4D-Var technic as seen on Fig. 5 , that displays ground truth, noisy image observation, image estimation at the end of the studied intervall.

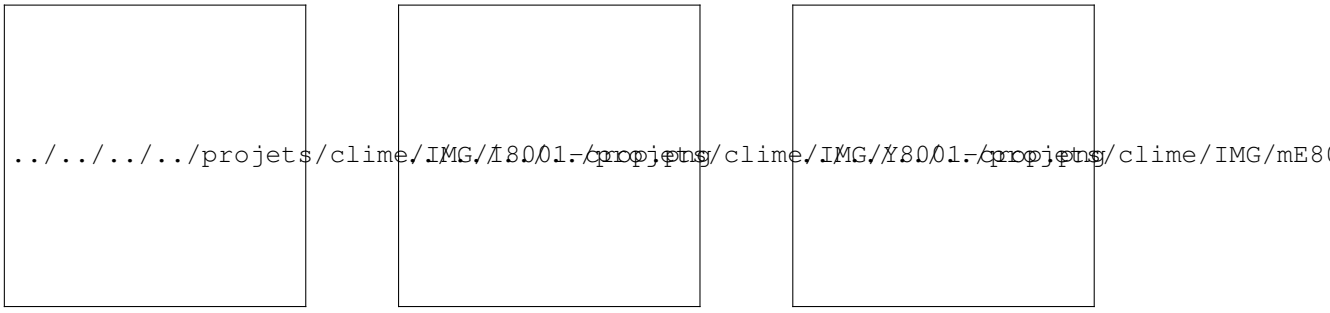


Figure 5. From left to right: Ground truth, image observation, result.

6.9. Fire application

6.9.1. Model evaluation for fire propagation

Participants: Vivien Mallet, Jean-Baptiste Fillipi [CNRS], Bahaa Nader [University of Corsica].

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, but currently, simulation errors are not sufficiently qualified and quantified.

We consider that the proper evaluation of a model requires to apply it to a large number of fires – instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number of fires (from the Prométhée database, http://www.promethee.com/) with the simulation model ForeFire (CNRS/University of Corsica) and evaluate the results with error measures [15]. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See illustration in Fig. 6 . We simulated 80 fires with four physical models, which proved that the most advanced models performed better overall, even though the input data is often inaccurate. We also carried out Monte Carlo simulations to evaluate the impact of the uncertainty in input data. We showed that the Monte Carlo approach led to a reliable forecasting system, which suggests that the probability densities derived from the simulations (see Fig. 6) may be useful information for preventive actions in an operational context.

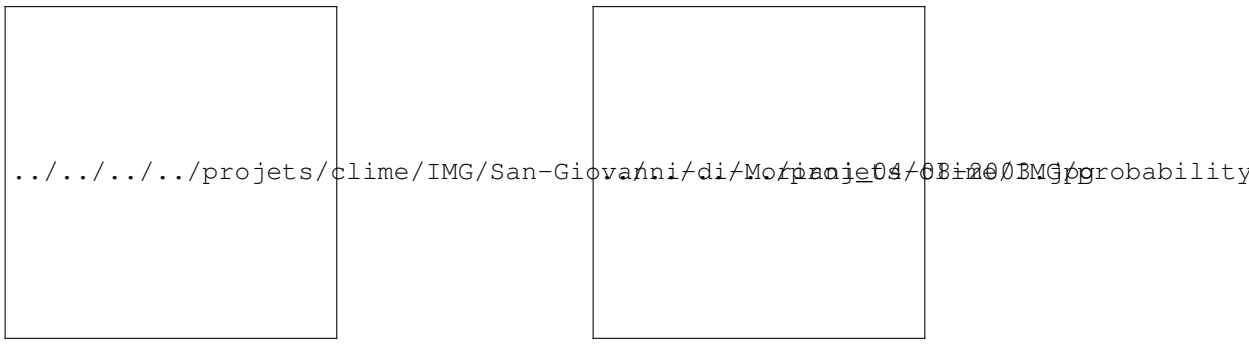


Figure 6. Left: Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica). Right: Burn probability as computed by a Monte Carlo simulation for a wildfire that was observed (red contour) in Corsica in 2003.

CONTRAINTES Project-Team

6. New Results

6.1. A Stronger Necessary Condition for the Multistationarity of Chemical Reaction Networks

Participant: Sylvain Soliman.

In the last thirty years, the conjecture of Thomas on the necessary presence of a positive circuit for the occurrence of multistationarity has opened a whole field of research, allowing better modeling and understanding of biochemical networks, especially in the emerging field of systems biology. However, if that aspect is striking in the field of discrete modeling of gene regulatory networks, it did not have the same impact in the Ordinary Differential Equations (ODE) based modeling community. This is mostly due to the fact that this necessary condition, the existence of a positive loop in the Jacobian of the ODE system, is almost always satisfied.

In [5] we improve on the ten years old proof by Soulé, using the structural information from the stoichiometric matrix of a biochemical reaction system. This allows us to state a more strict version of the famous Thomas' necessary condition for multistationarity. In particular, the obvious cases where Thomas' condition was trivially satisfied, mutual inhibition due to a multimolecular reaction and mutual activation due to a reversible reaction, can now easily be ruled out. The new condition makes it possible to use circuit analysis as an useful tool in the arsenal of the computational biologist, together with other structural methods.

6.2. Petri Net Analyses of Biochemical Reaction Networks using Constraint Logic Programming

Participants: François Fages, Thierry Martinez, Faten Nabli, Sylvain Soliman.

The Thesis of Faten Nabli [1] marks our achievements on the static analysis of biochemical reaction networks using Petri Net concepts and Constraint Logic Programming algorithms. This Thesis presents a Boolean model and two constraint-based methods for enumerating all minimal siphons and traps of a Petri net, by iterating the resolution of Boolean satisfiability problems executed with either a SAT solver or a CLP(B) program. The performances of these methods are compared with respect to a state-of-the-art algorithm from the Petri net community. On a benchmark with 80 Petri nets from the Petriweb database and 403 Petri nets from curated biological models of the Biomodels database, we show that miniSAT and CLP(B) solvers are overall both faster by two orders of magnitude with respect to the dedicated algorithm. Furthermore, we analyse why these programs perform so well on even very large biological models and show a polynomial time complexity result for Petri nets of fixed treewidth, using a similar theorem for constraint satisfaction problems with bounded treewidth constraint graphs. Faten Nabli has been hired with a Post-Doc position at Sanofi Paris.

6.3. Structural Model Reduction: CLP and SAT Solvers for Computing Subgraph Epimorphisms

Participants: François Fages, Steven Gay, Thierry Martinez, Francesco Santini, Sylvain Soliman.

This year, in [8], we have developed and compared CLP and SAT solvers on the NP-complete problem of deciding the existence of a subgraph epimorphism between two graphs. Our interest in this variant of graph matching problem stems from the study of model reductions in systems biology, where large systems of biochemical reactions can be naturally represented by bipartite digraphs of species and reactions. In this setting, model reduction can be formalized as the existence of a sequence of vertices, species or reaction, deletion and merge operations which transforms a first reaction graph into a second graph [2]. This problem is in turn equivalent to the existence of a subgraph (corresponding to delete operations) epimorphism (i.e. surjective homomorphism, corresponding to merge operations) from the first graph to the second. We show how subgraph epimorphism problems can be modeled as Boolean constraint satisfaction problems, and we compare CLP and SAT solvers on a large benchmark of reaction graphs from systems biology.

6.4. Quantitative Model Reduction: a CLP Solver for Computing Tropical Equilibrations

Participants: François Fages, Sylvain Soliman.

Model reduction is a central topic in computational systems biology and dynamical systems theory, for reducing the complexity of quantitative models, finding important parameters, and developing multi-scale models for instance. While perturbation theory is a standard mathematical tool to analyze the different time scales of a dynamical system, and decompose the system accordingly, tropical methods provide a simple algebraic framework to perform these analyses systematically in polynomial systems. The crux of these tropicalization methods is in the computation of tropical equilibrations. In [13], we show that constraint-based methods, using reified constraints for expressing the equilibration conditions, make it possible to numerically solve non-linear tropical equilibration problems, out of reach of standard computation methods. We illustrate this approach first with the reduction of simple biochemical mechanisms such as the Michaelis-Menten and Goldbeter-Koshland models, and second, with performance figures obtained on a large scale on the model repository `biomodels.net`.

This work is done in collaboration with Ovidiu Radulescu, Univ. Montpellier, in the context of a larger project about symbolic methods in systems biology with François Boullier, LIFL and Andras Weber, Univ. Bonn.

6.5. Species Minimization in Biochemical Reaction Computing

Participants: Hui-Ju Chiang, François Fages.

Engineering biochemical reactions for computational purposes is a common pursue in synthetic biology. In such design tasks, molecular species have to be carefully engineered to ensure modularity and orthogonality, and are scarce resources. Minimizing the number of involved molecular species is crucial to accomplish a complex computation within a confined biochemical environment.

In [10], we investigate an approach to species minimization by reusing modular and regular reactions in an asynchronous time-multiplexed fashion. Our method enhances not only species utility, but also re-programmability and robustness in realizing various logic circuits. A case study demonstrates the ease of design in realizing general logic computation, and simulation confirms the feasibility and robustness of the proposed method.

This work is done in collaboration with Jie-Hong Jiang and Katherine Chiang from NTU Taiwan in the context of a common project about biochemical programming.

6.6. Hybrid Composition and Simulation of Heterogeneous Biochemical Models

Participants: Hui-Ju Chiang, François Fages, Sylvain Soliman.

[2] Steven Gay, Sylvain Soliman, François Fages. A Graphical Method for Reducing and Relating Models in Systems Biology. Bioinformatics, 26(18):i575–i581, 2010.

Models of biochemical systems presented as a set of formal reaction rules with kinetic expressions can be interpreted with different semantics: as either deterministic Ordinary Differential Equations, stochastic continuous-time Markov Chains, Petri nets or Boolean transition systems. While the formal composition of reaction models can be syntactically defined as the (multiset) union of the reactions, the hybrid composition of models in different formalisms is a largely open issue.

In [7], we show that the combination of reaction rules with conditional events, as the ones already present in SBML, does provide the expressive power of hybrid automata and can be used in a non standard way to give meaning to the hybrid composition of heterogeneous models of biochemical processes. In particular, we show how hybrid differential-stochastic and hybrid differential-Boolean models can be compiled and simulated in this framework, through the specification of a high-level interface for composing heterogeneous models. This is illustrated by a hybrid stochastic-differential model of bacteriophage T7 infection, and by a reconstruction of the hybrid model of the mammalian cell cycle regulation of Singhania et al. as the composition of a Boolean model of cell cycle phase transitions and a differential model of cyclin activation.

6.7. Composition and Abstraction of Logical Influence Networks: Application to Multi-Cellular Systems

Participant: Grégory Batt.

Logical (Boolean or multi-valued) modelling is widely employed to study regulatory or signalling networks. Even though these discrete models constitute a coarse, yet useful, abstraction of reality, the analysis of large networks faces a classical combinatorial problem. In [4], we proposed to take advantage of the intrinsic modularity of inter-cellular networks to set up a compositional procedure that enables a significant reduction of the dynamics, yet preserving the reachability of stable states. To that end, we relied on process algebras, a well-established computational technique for the specification and verification of interacting systems.

We developed a novel compositional approach to support the logical modelling of interconnected cellular networks. First, we formalised the concept of logical regulatory modules and their composition. Then, we made this framework operational by transposing the composition of logical modules into a process algebra framework. Importantly, the combination of incremental composition, abstraction and minimisation using an appropriate equivalence relation (here the safety equivalence) yields huge reductions of the dynamics. We illustrated the potential of this approach with two case-studies: the Segment-Polarity and the Delta-Notch modules.

6.8. Identification of Biological Models from Single Cell Data: a Comparison between Mixed-Effects and Moment-based Inference

Participants: Grégory Batt, Andres Mauricio Gonzalez Vargas, Pascal Hersen, Artémis Llamosi, Jannis Uhlendorf.

Experimental techniques in biology such as microfluidic devices and time-lapse microscopy allow tracking of the gene expression in single cells over time. So far, few attempts have been made to fully exploit these data for modeling the dynamics of biological networks in cell populations. In [9], we compare two modeling approaches capable to describe cell-to-cell variability: Mixed-Effects (ME) models and the Chemical Master Equation (CME). We discuss how network parameters can be identified from experimental data and use real data of the HOG pathway in yeast to assess model quality.

For CME we rely on the identification approach proposed by Zechner et al. (PNAS, 2012), based on moments of the probability distribution involved in the CME. ME and moment-based (MB) inference will be also contrasted in terms of general features and possible uses in biology.

6.9. STL-based Analysis of TRAIL-induced Apoptosis Challenges the Notion of Type I/Type II Cell Line Classification

Participants: Grégory Batt, François Bertaux, Szymon Stoma.

Extrinsic apoptosis is a programmed cell death triggered by external ligands, such as the TNF-related apoptosis inducing ligand (TRAIL). Depending on the cell line, the specific molecular mechanisms leading to cell death may significantly differ. Precise characterization of these differences is crucial for understanding and exploiting extrinsic apoptosis. Cells show distinct behaviors on several aspects of apoptosis, including (i) the relative order of caspases activation, (ii) the necessity of mitochondria outer membrane permeabilization (MOMP) for effector caspase activation, and (iii) the survival of cell lines overexpressing Bcl2. These differences are attributed to the activation of one of two pathways, leading to classification of cell lines into two groups: type I and type II.

In [6] we challenge this type I/type II cell line classification. We encode the three aforementioned distinguishing behaviors in a formal language, called signal temporal logic (STL), and use it to extensively test the validity of a previously-proposed model of TRAIL-induced apoptosis with respect to experimental observations made on different cell lines. After having solved a few inconsistencies using STL-guided parameter search, we show that these three criteria do not define consistent cell line classifications in type I or type II, and suggest mutants that are predicted to exhibit ambivalent behaviors. In particular, this finding sheds light on the role of a feedback loop between caspases, and reconciliates two apparently-conflicting views regarding the importance of either upstream or downstream processes for cell-type determination. More generally, our work suggests that these three distinguishing behaviors should be merely considered as type I/II features rather than cell-type defining criteria. On the methodological side, this work illustrates the biological relevance of STL-diagrams, STL population data, and STL-guided parameter search implemented in the tool Breach. Such tools are well-adapted to the ever-increasing availability of heterogeneous knowledge on complex signal transduction pathways.

6.10. Single Cell Models and Models of Populations: A Mixed Effect Approach

Participants: Grégory Batt, Andres Mauricio Gonzalez Vargas, Pascal Hersen, Artémis Llamosi.

For a long time, experiments and models of gene expression were mainly based on the mean behavior of a population of cells. Although observed early, it is only recently that experimental technique allowed detailed investigation of variability in this process. Since the pioneering work of Elowitz and colleagues, a distinction is drawn between what is called intrinsic and extrinsic variability or noise. Intrinsic noise originates in the randomness of chemical reactions within a cell whether extrinsic noise is the variation in between cells at a given time. Extrinsic variability is associated with population heterogeneity in the concentrations of ribosomes or other molecular players or processes relevant to gene expression (RNAPolII concentration, degradation and dilution rates etc.).

In this work, we propose a modelling framework for gene expression based on a system of ODEs with random parameters following a distribution across the population of cells. In this context, each cell has its own identity which is represented by the value of its parameters. With this model we ask how much of the long term variability can be explained by extrinsic variability alone. We produced long term, time lapse and single-cell data of repeated gene induction in Saccharomyces cerevisiae. One experiment was treated as learning set whereas two were used as test sets. From the learning set, we are able to infer single cell parameters and population distributions which represent accurately in terms of mean and variance the variability in the population. These learned population distributions allowed good predictions on both the learning and test sets.

Our study demonstrates also that the way inference of single cell parameters and distributions is performed is crucial to achieve good performance. Best results being found by joint estimation of the parameters for single cells and for the whole population. With this technique, we noted that very decent fits of the population dynamics can be obtained by estimating only on a very limited number of cells. Concerning the quality of single cell parameters inferred, we validated the presence of an expected significant correlation between the dilution rate and the measured single cell growth rate. This motivates the use of this tool in order to investigate the origins of extrinsic noise, by correlating single cell parameters with measured candidate factors of gene expression variability such as cell density, cell size or age.

6.11. Coupled Model of the Cell Cycle and Circadian Clock

Participants: François Fages, Sylvain Soliman, Denis Thieffry, Pauline Traynard.

Recent advances in cancer chronotherapy techniques support the evidence that there exist important links between the cell cycle and the circadian clock genes. One purpose for modeling these links is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characterictics. This is at the heart of our participation in collaboration with the EPI BANG in the EraNet SysBio project C5Sys, follow up of the former EU STREP project TEMPO.

This year we have pursued the investigation of the effect of transcription inhibition during mitosis, as a reverse coupling from the cell cycle to the circadian clock. We use quantitative temporal logic constraints and the parallel version of BIOCHAM for parameter search, running on the Jade cluster of 10000 processors at the GENCI CINES, to couple dynamical models in high dimension and fit models to experimental data time series obtained in Franck Delaunay's lab in Nice, CNRS. We are defining a series of common temporal logic patterns and *ad hoc* schemes for computing their validity domain on a given trace, more efficiently than by the generic method implemented in BIOCHAM.

6.12. Solving Mixed Shapes Packing Problems by Continuous Optimization with the CMA Evolution Strategy

Participants: François Fages, Thierry Martinez, Lumadaiara Do Nascimento Vitorino.

Bin packing is a classical combinatorial optimization problem which has a wide range of real-world applications in industry, logistics, transport, parallel computing, circuit design and other domains. While usually presented as discrete problems, in [12] we consider continuous packing problems including curve shapes, and model these problems as continuous optimization problems with a multi-objective function combining non-overlapping with minimum bin size constraints. More specifically, we consider the covariance matrix adaptation evolution strategy (CMA-ES) with a nonoverlapping and minimum size objective function in either two or three dimensions. Instead of taking the intersection area as measure of overlap, we propose other measures, monotonic with respect to the intersection area, to better guide the search. In order to compare this approach to previous work on bin packing, we first evaluate CMA-ES on Korf's benchmark of consecutive sizes square packing problems, for which optimal solutions are known, and on a benchmark of circle packing problems. We show that on square packing, CMA-ES computes solutions at typically 14% of the optimal cost, with the time limit given by the best dedicated algorithm for computing optimal solutions, and that on circle packing, the computed solutions are at 2% of the best known solutions. We then consider generalizations of this benchmark to mixed squares and circles, boxes, spheres and cylinders packing problems, and study a real-world problem for loading boxes and cylinders in containers. These hard problems illustrate the interesting trade-off between generality and efficiency in this approach.

6.13. Railway Time Tabling Optimization with CMA-ES and Greedy Heuristics

Participants: François Fages, David Fournier, Thierry Martinez, Sylvain Soliman.

The problem of reducing energy consumption in public transportation has received increasing attention over the last years. Most metros have energy regenerative braking systems, which allow them to produce electric energy when they brake. We study the problem of optimizing the energy consumption of a metro line by modifying the timetable, in order to maximize the actual reuse of the regenerative energy. This is achieved by synchronizing the braking and acceleration phases of the metros, through slight modifications of the stopping times in stations. In an article in preparation, we present a constraint-based model of the electric network of the line, which is used to evaluate the energy consumption at each instant, and to compute a distribution matrix for approximating the potential energy transfers between metros. The optimization of the timetable is then performed by an evolutionary algorithm using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES from Nikolaus Hansen, EPI TAO). On real data, this strategy shows energy savings ranging from 2.38% to 4.54%. Furthermore, these savings are shown to be robust with respect to perturbations of the dwell times.

CRYPT Team (section vide)

DEDUCTEAM Exploratory Action

6. New Results

6.1. Dedukti

The version 2.0 of the Dedukti system, developed by Ronan Saillard, has been released in July 2013. It is based on an improved version of the $\lambda\Pi$-calculus modulo where rewrite rules are explicitly added [31], and where the conditions for typing the rewrite rules are weakened.

This version is fully written in OCaml. It is smaller, far more efficient than the previous version, and permits to type-check much bigger files.

New features include a better reporting of errors, an interactive mode, an export functionality from Dedukti to the MMT format [53], and non-linear pattern matching.

6.2. Embeddings in the $\lambda\Pi$-calculus modulo

A new version of Coqine has been developed by Ali Assaf. This version is designed using a Coq plugin architecture, which allows for a smoother integration with Coq's code base and alleviates problems of maintainability that affected the previous version.

The implementation of Holide has been improved, by Ali Assaf. This improved version incorporates sharing at the level of terms and types. This optimization allows to reduce the type-checking time of the OpenTheory standard library from more than 30 minutes to less than 1 minute.

Catherine Dubois and Raphaël Cauderlier have studied a translation in the $\lambda\Pi$-calculus modulo of features coming from object oriented programming languages, such as inheritance and late binding. This compilation scheme has been applied to produce a new back-end for FoCaLize called Focalide, through a compilation to Dedukti. This translation can benefit from the flexibility of Dedukti to deal with more dynamic object-oriented languages than FoCaLiZe; they are currently working on generalizing this translation using ζ-calculus as a theoretical foundation for objects.

Resolution and superposition are proof-search methods that are used in state-of-the-art first-order automated theorem provers such as iProver, Vampire, E or SPASS. A shallow embedding of resolution and superposition proofs in the $\lambda\Pi$-calculus modulo has been proposed by Guillaume Burel, thus offering a way to check these proofs in a trusted setting, and to combine them with other proofs. This embedding has been implemented in particular as a backend of iProver Modulo, therefore allowing to check proofs found by iProver Modulo using Dedukti [20].

A shallow embedding in Dedukti of the tableaux proofs generated by Zenon modulo has been designed and implemented by Frédéric Gilbert [22], [23]. The embedding is based on a refined version of previous double-negation translations, introducing as less as possible double negations. This optimization has shown that more than half of the proofs found by Zenon modulo are not using the excluded-middle law, therefore being purely intuitionistic.

6.3. Automated Theorem Proving

Mélanie Jacquel (*Siemens*) and David Delahaye developed *Super Zenon* [5], a generalization of the extension of *Zenon* to superdeduction to handle any first order theory. To do so, they designed heuristics able to automatically transform axioms of a theory into rewrite rules. This new tool has been tested over the first order problems of the TPTP library and a significant increase has been observed. A first distribution of this tool (under GPL licence) is planned in the first months of 2014. In addition, an integration to the *Rodin* platform is also planned with the help of Laurent Voisin (*Systerel*). This integration should allow us to apply this tool in the context of *Event-B*.

Pierre Halmagrand, David Delahaye, Damien Doligez, and Olivier Hermant developed *Zenon Modulo* [22], [23], an extension of *Zenon* to Deduction modulo. Like *Super Zenon*, this new tool is able to deal with any first order theory and relies on an heuristic able to automatically transform axioms of a theory into rewrite rules. This tool has also been tested over the first order problems of the TPTP library and a similar increase of performance (compared to *Super Zenon*) has been observed. Frédéric Gilbert has developed a *Dedukti* backend for this extension, which is based on a double-negation transformation that allows us to transform classical proofs produced by *Zenon Modulo* into intuitionistic proofs in *Dedukti*. This tool is intended to be applied in the framework of the *BWare* project in order to automatically verify proof obligations coming from the modeling of industrial applications. To do so, the idea is to manually transform the *B* set theory into a theory modulo and provide it to *Zenon Modulo* in order to verify the proof obligations of the *BWare* project.

Guillaume Burel and Simon Cruanes have designed a method to scan sets of first-order clauses in order to detect the presence of instances of axiomatic theories (group structures, total orderings, etc.), even during a saturation process (so that theories that only become apparent during the proof search can be detected) [21]. To this end, they introduced the concept of *meta-prover*, a Datalog system that reasons over properties of the problem, and communicates with the saturation prover. This technique made some applications possible, such as the use of generic lemma and an equational redundancy criterion for some theories, and was implemented in Zipperposition.

Simon Cruanes has been working on superposition modulo linear arithmetic, using Zipperposition as a test bed. The focus is on problems with rational or integer arithmetic mixed with first-order reasoning, an area in which SMT solvers struggle. The work is still preliminary, but shows promising results.

Depending on the logic for finite structures, which is defined by Gilles Dowek and Ying Jiang (Beijing), Kailiang Ji has extended the use of proof search algorithms in Deduction modulo to automatically prove some graph properties, such as (un)reachability, which can be described by CTL formulas. A technical report about this has been given on Locali 2013 in Beijing.

Together with Tayssir Touili (University Paris Diderot) Hugo Macedo has shown how to advance the performance of the application of model checking techniques in the domain of malicious software detection. The work consisted in leveraging the reachability analysis used in the model checking of pushdown systems to infer malicious behavior patterns from known malware. From such new application a malware detection tool was prototyped and put to the test with instances of "in the wild" (real world) malicious software. This work was published in a large security venue and the details about the technique follow in [29].

Kim-Quyen Ly extended her formally-proved (in Coq) automated termination-certificate (for first-order rewrite systems) verifier Rainbow for dealing with certificates using arguments filtering [39] and other termination techniques.

6.4. Proof theory

The conservativity of the embedding of pure type systems in the $\lambda\Pi$-calculus modulo was proved by Ali Assaf. This result extends those of Cousineau and Dowek [46] and further justifies the use of the $\lambda\Pi$-calculus modulo as a logical framework. This embedding is the basis for the automated translation tools Holide and Coqine.

Frédéric Blanqui, Jean-Pierre Jouannaud (Univ. Paris 11) and Albert Rubio (Technical University of Catalonia) have developed a method aiming at carrying out termination proofs for higher-order calculi. CPO appears to be the ultimate improvement of the higher-order recursive path ordering (HORPO) [45] in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard, therefore explaining the name of the improved ordering. It has been shown that CPO allows to consider higher-order rewrite rules in a simple type discipline with inductive types, that most of the guards present in the recursive calls of its core definition cannot be relaxed in any natural way without losing well-foundedness, and that the precedence on function symbols cannot be made more liberal anymore.

Frédéric Blanqui worked on the formalization in the Coq proof assistant of various definitions of the notion of α-equivalence on pure λ-terms. In particular, he formalized and formally proved equivalent the definitions

of Church (1932), Curry and Feys (1958), Krivine (1993), and Gabbay and Pitts (1999). This work is freely available from the CoLoR library released on December 13th.

Alejandro Díaz-Caro and Gilles Dowek have introduced an extension of λ-calculus with pairs where isomorphic types are equated. Identifying some types requires to also identify some terms via an equivalence relation on terms, leading to an interesting calculus, which is related to several known non-deterministic and probabilistic calculi. A preliminary version of this work has been published on [24]. A complete version in simple types, with its proof of normalisation, is currently under review.

Together with Ying Jiang, Gilles Dowek has started to investigate the links between model-checking and proof-checking. This has materialized by an encoding of CTL for a finite model in predicate logic and by the definition of a proof-system for CTL.

Olivier Hermant has studied optimized versions of double-negation translations, that allow to switch between classical and intuitionistic logics. Such an algorithm has been implemented in Zenon's backend to Dedukti by Frédéric Gilbert. Gilles Dowek has given new version of Gödel's translation of classical logic into constructive logic. This translation is homomorphic, hence it can be seen as a mere definition of the classical connectives from the constructive ones.

6.5. Safety of aerospace systems

Pierre Néron has designed a method to transform straight line programs, such as those used in some aerospace systems into others that do not use some operations such as, square roots and divisions that cannot be performed exactly on decimal numbers. To this end he has defined a new notion of anti-unification, called *constrained anti-unification*, and a new anti-unification algorithm.

6.6. Models of Computation

Alejandro Díaz-Caro and Gilles Dowek have shown how to provide a structure of probability space to the set of execution traces on a non-confluent abstract rewrite system, by defining a variant of a Lebesgue measure on the space of traces. Then, they showed how to use this probability space to transform a non-deterministic calculus into a probabilistic one. As an example, they applied this technique to the previously introduced non-deterministic calculus. [25]

Ali Assaf and Alejandro Díaz-Caro, together with Simon Perdrix (Nancy), Christine Tasson (PPS) and Benoît Valiron (PPS) have determined the relationship between the algebraic λ-calculus, a fragment of the differential λ-calculus and the linear-algebraic λ-calculus, a candidate λ-calculus for quantum computation. Both calculi are algebraic: each one is equipped with an additive and a scalar-multiplicative structure, and their set of terms is closed under linear combinations. However, the two languages were built using different approaches: the former is a call-by-name language whereas the latter is call-by-value; the former considers algebraic equalities whereas the latter approaches them through rewrite rules. They have analysed how these different approaches relate to one another, proposing four canonical languages based on each of the possible choices: call-by-name versus call-by-value, algebraic equality versus algebraic rewriting. They have shown that the various languages simulate one another. Preliminary versions of this work where published in [47] and [41]. Now they are working on a journal version filling the gaps between these previous works.

Together with Pablo Arrighi (Grenoble) and Benoît Valiron (PPS), Alejandro Díaz-Caro has described a type system for the linear-algebraic lambda-calculus. The type system accounts for the linear-algebraic aspects of this extension of lambda-calculus: It is able to statically describe the linear combinations of terms that will be obtained when reducing the programs. This gives rise to an original type theory where types, in the same way as terms, can be superposed into linear combinations. They have proven that the resulting typed lambda-calculus is strongly normalising and features a weak subject reduction. In addition, they have shown how to naturally encode matrices and vectors in this typed calculus [34].

Gilles Dowek has investigated a new definition of the notion of a chaotic system that can be applied to discrete systems and that is compatible with the principle of a finite density of information.

The paper Call-by-value non-determinism in a linear logic type discipline by Alejandro Díaz-Caro, Giulio Manzonetto and Michele Pagani has been published [26].

The paper Universality in two dimensions of Gilles Dowek and Nachum Dershowitz has been published.

The paper Linear-algebraic lambda-calculus: higher-order, encodings and confluence of Pablo Arrighi and Gilles Dowek has been published.

The book *Lambda Calculus with Types* , written by Henk Barendregt, Wil Dekkers, Richard Statman, and 11 contributors, including Gilles Dowek, has been published.

6.7. Constraint solving

Catherine Dubois has extended the formally verified constraint solver (on finite domains) she has developed with Matthieu Carlier and Arnaud Gotlieb with a new local consistency property (bound-consistency).

DYOGENE Project-Team

6. New Results

6.1. Ancillary service to the grid from deferrable loads: the case for intelligent pool pumps in Florida

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. In [28], we focus on pool pumps, and how they can be used to provide ancillary service to the grid for maintaining demand-supply balance. A Markovian Decision Process (MDP) model is introduced for an individual pool pump. 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 pools is then developed by examining the mean field limit. A key innovation is an 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. Simulations are provided to illustrate the accuracy of the approximations and effectiveness of the proposed control approach.

6.2. Impact of Storage on the Efficiency and Prices in Real-Time Electricity Markets

In [19] we study the effect of energy-storage systems in dynamic real-time electricity markets. We consider that demand and renewable generation are stochastic, that real-time production is affected by ramping constraints, and that market players seek to selfishly maximize their profit. We distinguish three scenarios, depending on the owner of the storage system: (A) the supplier, (B) the consumer, or (C) a stand-alone player. In all cases, we show the existence of a competitive equilibrium when players are price-takers (they do not affect market prices). We further establish that under the equilibrium price process, players' selfish responses coincide with the social welfare-maximizing policy computed by a (hypothetical) social planner. We show that with storage the resulting price process is smoother than without. We determine empirically the storage parameters that maximize the players' revenue in the market. In the case of consumer-owned storage, or a stand-alone storage operator (scenarios B and C), we find that they do not match socially optimal parameters. We conclude that consumers and the stand-alone storage operator (but not suppliers) have an incentive to under-dimension their storage system. In addition, we determine the scaling laws of optimal storage parameters as a function of the volatility of demand and renewables. We show, in particular, that the optimal storage energy capacity scales as the volatility to the fourth power.

6.3. Risk-Aware SLA Negotiation

In order to assure Quality of Service (QoS) connectivity, Network Service Providers (NSPs) negotiate Service Level Agreements (SLAs). However, a committed SLA might fail to respect its QoS promises. In such a case, the customer is refunded. To maximize their revenues, the NSPs must deal with risks of SLA violations, which are correlated to their network capacities. Due to the complexity of the problem, we first study in [21], a system with one NSP provider and give a method to compute its risk-aware optimal strategy using (max; +)-algebras. Using the same method, we study the case where two NSPs collaborate and the case where they compete, and we derive the Price of Anarchy. This method provides optimal negotiation strategies but, when modeling customers' reaction to SLA failure, analytical results do not hold. Hence, we propose a learning framework that chooses the NSP risk-aware optimal strategy under failures capturing the impact of reputation. Finally, by simulation, we observe how the NSP can benefit from such a framework.

6.4. Impact of Rare Alarms on Event Correlation

Nowadays, telecommunication systems are growing more and more complex, generating a large amount of alarms that cannot be effectively managed by human operators. The problem is to detect significant combinations of alarms describing an issue in real-time. In [18], we present a powerful heuristic algorithm that constructs dependency graphs of alarm patterns. More precisely, it highlights patterns extracted from an alarm flow obtained from a learning process with a small footprint on network management system performance. This algorithm helps to detect issues in real-time by effectively delivering concise alarm patterns. Furthermore, it allows the proactive analysis of the functioning of a network by computing the general trends of this network. We evaluate our algorithm on an optical network alarm data set of an existing operator. We find similar results as the expert analysis performed for this operator by Alcatel-Lucent Customer Services.

6.5. Some Synchronization Issues in OSPF Routing

A routing protocol such as OSPF has a cyclic behavior to regularly update its view of the network topology. Its behavior is divided into periods. Each period produces a flood of network information messages. We observe a regular activity in terms of messages exchanges and filling of receive buffers in routers. [17] examines the consequences of possible overlap of activity between periods, leading to a buffer overflow. OSPF allows “out of sync" flows by considering an initial delay (phase). We study the optimum calculation of these offsets to reduce the load, while maintaining a short period to ensure a protocol reactive to topology changes. Such studies are conducted using a simulated Petri net model. A heuristic for determining initial delays is proposed. A core network in Germany serves as illustration.

6.6. Exact Worst-case Delay in FIFO-multiplexing Feed-forward Networks

In this paper 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.

6.7. Fast weak–KAM integrators for separable Hamiltonian systems

We consider a numerical scheme for Hamilton–Jacobi equations based on a direct discretization of the Lax–Oleinik semi–group. We prove that this method is convergent with respect to the time and space stepsizes provided the solution is Lipschitz, and give an error estimate. Moreover, we prove that the numerical scheme is a *geometric integrator* satisfying a discrete weak–KAM theorem which allows to control its long time behavior. Taking advantage of a fast algorithm for computing min–plus convolutions based on the decomposition of the function into concave and convex parts, we show that the numerical scheme can be implemented in a very efficient way.

6.8. Probabilistic cellular automata, invariant measures, and perfect sampling

A probabilistic cellular automaton (PCA) can be viewed as a Markov chain. The cells are updated synchronously and independently, according to a distribution depending on a finite neighborhood. In [9], we investigate the ergodicity of this Markov chain. A classical cellular automaton is a particular case of PCA. For a one-dimensional cellular automaton, we prove that ergodicity is equivalent to nilpotency, and is therefore undecidable. We then propose an efficient perfect sampling algorithm for the invariant measure of an ergodic PCA. Our algorithm does not assume any monotonicity property of the local rule. It is based on a bounding process which is shown to also be a PCA. Last, we focus on the PCA majority, whose asymptotic behavior is unknown, and perform numerical experiments using the perfect sampling procedure.

6.9. Density Classification on Infinite Lattices and Trees

Consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter p. In [7], we address the density classification problem, that is, we want to design a (probabilistic or deterministic)cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether p is smaller or larger than 1/2. Precisely, the trajectories should converge to the uniform configuration with only 0's if p<1/2, and only 1's if p>1/2. We present solutions to the problem on the regular grids of dimension d, for any d>1, and on the regular infinite trees. For the bi-infinite line, we propose some candidates that we back up with numerical simulations.

6.10. Semi-infinite paths of the radial spanning tree

In the paper [4], in collaboration with David Coupier and Viet Chi Tran of Lille 1, we study the semi-infinite paths of the radial spanning tree (RST) of a Poisson point process in the plane using Stochastic Geometry. We first show that the expectation of the number of intersection points between semi-infinite paths and the sphere with radius r grows sublinearly with r. Then, we prove that in each (deterministic) direction, there exists with probability one a unique semi-infinite path, framed by an infinite number of other semi-infinite paths of close asymptotic directions. The set of (random) directions in which there are more than one semi-infinite paths is dense in $[0, 2\pi)$. It corresponds to possible asymptotic directions of competition interfaces. We show that the RST can be decomposed in at most five infinite subtrees directly connected to the root. The interfaces separating these subtrees are studied and simulations are provided.

6.11. Generating functionals of random packing point processes

In the paper [45], we study the generating functionals of a class of random packing point processes of the Matérn type. Consider a symmetrical conflict relationship between the points of a point process. The Matérn type constructions provide a generic way of selecting a subset of this point process which is conflict-free. The simplest one consists in keeping only conflict-free points. There is however a wide class of Matérn type processes based on more elaborate selection rules and providing larger sets of selected points. The general idea being that if a point is discarded because of a given conflict, there is no need to discard other points with which it is also in conflict. The ultimate selection rule within this class is the so called Random Sequential Adsorption, where the cardinality of the sequence of conflicts allowing one to decide whether a given point is selected is *not* bounded. The present paper provides a sufficient condition on the span of the conflict relationship under which all the above point processes are well defined when the initial point process is Poisson. It then establishes, still in the Poisson case, a set of differential equations satisfied by the probability generating functionals of these Matérn type point processes. Integral equations are also given for the Palm distributions.

6.12. Clustering and percolation of point processes

We are interested in phase transitions in certain percolation models on point processes and their dependence on clustering properties of the point processes. In [5], we show that point processes with smaller void probabilities and factorial moment measures than the stationary Poisson point process exhibit non-trivial phase transition in the percolation of some coverage models based on level-sets of additive functionals of the point process. Examples of such point processes are determinantal point processes, some perturbed lattices, and more generally, negatively associated point processes. Examples of such coverage models are k-coverage in the Boolean model (coverage by at least k grains) and SINR-coverage (coverage if the signal-to-interference-and-noise ratio is large). In particular, we answer in affirmative the hypothesis of existence of phase transition in the percolation of k-faces in the C̆ech simplicial complex (also called clique percolation) on point processes which cluster less than the Poisson process. We also construct a Cox point process, which is "more clustered" than the Poisson point process and whose Boolean model percolates for arbitrarily small radius. This shows that clustering (at least, as detected by our specific tools) does not always "worsen" percolation, as well as that upper-bounding this clustering by a Poisson process is a necessary assumption for the phase transition to hold.

6.13. Using Poisson processes to model lattice cellular networks

An almost ubiquitous assumption made in the stochastic-analytic approach to study of the quality of user-service in cellular networks is Poisson distribution of base stations, often completed by some specific assumption regarding the distribution of the fading (e.g. Rayleigh). The former (Poisson) assumption is usually (vaguely) justified in the context of cellular networks, by various irregularities in the real placement of base stations, which ideally should form a lattice (e.g. hexagonal) pattern. In the first part of [14] we provide a different and rigorous argument justifying the Poisson assumption under sufficiently strong log-normal shadowing observed in the network, in the evaluation of a natural class of the typical-user service-characteristics (including path-loss, interference, signal-to-interference ratio, spectral efficiency). Namely, we present a Poisson-convergence result for a broad range of stationary (including lattice) networks subject to log-normal shadowing of increasing variance. We show also for the Poisson model that the distribution of all these typical-user service characteristics does not depend on the particular form of the additional fading distribution. Our approach involves a mapping of 2D network model to 1D image of it "perceived" by the typical user. For this image we prove our Poisson convergence result and the invariance of the Poisson limit with respect to the distribution of the additional shadowing or fading. Moreover, in the second part of the paper we present some new results for Poisson model allowing one to calculate the distribution function of the SINR in its whole domain. We use them to study and optimize the mean energy efficiency in cellular networks.

6.14. Compactification of the Action of a Point-Shift on the Palm Probability of a Point Process

In collaboration with Mir-Omid Haji-Mirsadeghi (Sharif University, Iran) [50], we analyzed the compactification of Palm probabilities by the action of a point–shift. A point-shift maps, in a translation invariant way, each point of a stationary point process Φ to some point of Φ. The initial motivation of this paper is the construction of probability measures, defined on the space of counting measures with an atom at the origin, which are left invariant by a given point-shift f. The point-shift probabilities of Φ are defined from the action of the semigroup of point-shift translations on the space of Palm probabilities, and more precisely from the compactification of the orbits of this semigroup action. If the point-shift probability is uniquely defined, and if f is continuous with respect to the vague topology, then the point-shift probability of Φ provides a solution to the initial question. Point-shift probabilities are shown to be a strict generalization of Palm probabilities: when the considered point-shift f is bijective, the point-shift probability of Φ boils down to the Palm probability of Φ. When it is not bijective, there exist cases where the point-shift probability of Φ is the law of Φ under the Palm probability of some stationary thinning Ψ of Φ. But there also exist cases where the point-shift probability of Φ is singular w.r.t. the Palm probability of Φ and where, in addition, it cannot be the law of Φ under the Palm probability of any stationary point process Ψ jointly stationary with Φ. The paper also gives a criterium of existence of the point-shift probabilities of a stationary point process and discusses uniqueness. The results are illustrated through several examples.

6.15. A Stochastic Geometry Framework for Analyzing Pairwise-Cooperative Cellular Networks

Cooperation in cellular networks has been recently suggested as a promising scheme to improve system performance, especially for cell-edge users. In [34], we use stochastic geometry to analyze cooperation models where the positions of Base Stations (BSs) follow a Poisson point process distribution and where Voronoi cells define the planar areas associated with them. For the service of each user, either one or two BSs are involved. If two, these cooperate by exchange of user data and channel related information with conferencing over some backhaul link. Our framework generally allows variable levels of channel information at the transmitters. In this paper we investigate the case of limited channel state information for cooperation (channel phase, second neighbour interference), but not the fully adaptive case which would require considerable feedback. The total per-user transmission power is further split between the two transmitters and a common message is encoded. The decision for a user to choose service with or without cooperation is directed by a family of

geometric policies depending on its relative position to its two closest base stations. An exact expression of the network coverage probability is derived. Numerical evaluation allows one to analyze significant coverage benefits compared to the non-cooperative case. As a conclusion, cooperation schemes can improve system performance without exploitation of extra network resources.

6.16. SINR-based k-coverage probability in cellular networks with arbitrary shadowing

In [20], we give numerically tractable, explicit integral expressions for the distribution of the signal-to-interference-and-noise-ratio (SINR) experienced by a typical user in the down-link channel from the k-th strongest base stations of a cellular network modelled by Poisson point process on the plane. Our signal propagation-loss model comprises of a power-law path-loss function with arbitrarily distributed shadowing, independent across all base stations, with and without Rayleigh fading. Our results are valid in the whole domain of SINR, in particular for $SINR < 1$, where one observes multiple coverage. In this latter aspect our paper complements previous studies reported in [55].

6.17. Equivalence and comparison of heterogeneous cellular networks

In [15], we consider a general heterogeneous network in which, besides general propagation effects (shadowing and/or fading), individual base stations can have different emitting powers and be subject to different parameters of Hata-like path-loss models (path-loss exponent and constant) due to, for example, varying antenna heights. We assume also that the stations may have varying parameters of, for example, the link layer performance (SINR threshold, etc). By studying the *propagation processes* of signals received by the typical user from all antennas marked by the corresponding antenna parameters, we show that seemingly different heterogeneous networks based on Poisson point processes can be equivalent from the point of view a typical user. These neworks can be replaced with a model where all the previously varying propagation parameters (including path-loss exponents) are set to constants while the only trade-off being the introduction of an isotropic base station density. This allows one to perform analytic comparisons of different network models via their isotropic representations. In the case of a constant path-loss exponent, the isotropic representation simplifies to a homogeneous modification of the constant intensity of the original network, thus generalizing a previous result showing that the propagation processes only depend on one moment of the emitted power and propagation effects. We give examples and applications to motivate these results and highlight an interesting observation regarding random path-loss exponents.

6.18. How user throughput depends on the traffic demand in large cellular networks: a typical cell analysis and real network measurements

In [40], we assume a space-time Poisson process of call arrivals on the infinite plane, independently marked by data volumes and served by a cellular network modeled by an infinite ergodic point process of base stations. Each point of this point process represents the location of a base station that applies a processor sharing policy to serve users arriving in its vicinity, modeled by the Voronoi cell, possibly perturbed by some random signal propagation effects. User service rates depend on their signal-to-interference-and-noise ratios with respect to the serving station. Little's that allows to express the mean user throughput in any region of this network model as the ratio of the mean traffic demand to the steady-state mean number of users in this region. Using ergodic arguments and the Palm theoretic formalism, we define a global mean user throughput in the cellular network and prove that it is equal to the ratio of mean traffic demand to the mean number of users in the steady state of the "typical cell" of the network. Here, both means account for double averaging: over time and network geometry, and can be related to the per-surface traffic demand, base-station density and the spatial distribution of the signal-to-interference-and-noise ratio. This latter accounts for network irregularities, shadowing and cell dependence via some cell-load equations. Inspired by the analysis of the typical cell, we propose also a simpler, approximate, but fully analytic approach, called the mean cell approach. The key quantity explicitly calculated in this approach is the cell load. In analogy to the load factor of the (classical) M/G/1 processor

sharing queue, it characterizes the stability condition, mean number of users and the mean user throughput. We validate our approach comparing analytical and simulation results for Poisson network model to real-network measurements.

6.19. Analysis of a Proportionally Fair and Locally Adaptive spatial Aloha in Poisson Networks

The proportionally fair sharing of the capacity of a Poisson network using Spatial-Aloha leads to closed-form performance expressions in two extreme cases: (1) the case without topology information, where the analysis boils down to a parametric optimization problem leveraging stochastic geometry; (2) the case with full network topology information, which was recently solved using shot-noise techniques. In [37], we show that there exists a continuum of adaptive controls between these two extremes, based on local stopping sets, which can also be analyzed in closed form. We also show that these control schemes are implementable, in contrast to the full information case which is not. As local information increases, the performance levels of these schemes are shown to get arbitrarily close to those of the full information scheme. The analytical results are combined with discrete event simulation to provide a detailed evaluation of the performance of this class of medium access controls.

6.20. Optimal Rate sampling in 802.11 Systems

In 802.11 systems, Rate Adaptation (RA) is a fundamental mechanism allowing transmitters to adapt the coding and modulation scheme as well as the MIMO transmission mode to the radio channel conditions, and in turn, to learn and track the (mode, rate) pair providing the highest throughput. So far, the design of RA mechanisms has been mainly driven by heuristics. In contrast, in [42], we rigorously formulate such design as an online stochastic optimisation problem. We solve this problem and present ORS (Optimal Rate Sampling), a family of (mode, rate) pair adaptation algorithms that provably learn as fast as it is possible the best pair for transmission. We study the performance of ORS algorithms in both stationary radio environments where the successful packet transmission probabilities at the various (mode, rate) pairs do not vary over time, and in non-stationary environments where these probabilities evolve. We show that under ORS algorithms, the throughput loss due to the need to explore sub-optimal (mode, rate) pairs does not depend on the number of available pairs, which is a crucial advantage as evolving 802.11 standards offer an increasingly large number of (mode, rate) pairs. We illustrate the efficiency of ORS algorithms (compared to the state-of-the-art algorithms) using simulations and traces extracted from 802.11 test-beds.

6.21. Flooding in Weighted Sparse Random Graphs

In [3], we study the impact of edge weights on distances in sparse random graphs. We interpret these weights as delays and take them as independent and identically distributed exponential random variables. We analyze the weighted flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node and the weighted diameter corresponding to the largest distance between any pair of vertices. Under some standard regularity conditions on the degree sequence of the random graph, we show that these quantities grow as the logarithm of n when the size of the graph n tends to infinity. We also derive the exact value for the prefactor. These results allow us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.

6.22. Viral Marketing On Configuration Model

In [38], we consider propagation of influence on a Configuration Model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset of its neighbours. Our (enhanced) model is described by the total degree of the typical vertex, representing the total number of its neighbours and the transmitter degree, representing the number of neighbours it is able to influence. We give a condition

involving the joint distribution of these two degrees, which if satisfied would allow with high probability the influence to reach a non-negligible fraction of the vertices, called a *big (influenced) component*, provided that the source vertex is chosen from a set of *good pioneers*. We show that asymptotically the big component is essentially the same, regardless of the good pioneer we choose, and we explicitly evaluate the asymptotic relative size of this component. Finally, under some additional technical assumption we calculate the relative size of the set of good pioneers. The main technical tool employed is the "fluid limit" analysis of the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed. This method was introduced in [59] to study the giant component of the configuration model. Using this approach we study also a reverse dynamic, which traces all the possible sources of influence of a given vertex, and which by a new "duality" relation allows to characterise the set of good pioneers.

6.23. Pioneers of Influence Propagation in Social Networks

With the growing importance of corporate viral marketing campaigns on online social networks, the interest in studies of influence propagation through networks is higher than ever. In a viral marketing campaign, a firm initially targets a small set of pioneers and hopes that they would influence a sizeable fraction of the population by diffusion of influence through the network. In general, any marketing campaign might fail to go viral in the first try. As such, it would be useful to have some guide to evaluate the effectiveness of the campaign and judge whether it is worthy of further resources, and in case the campaign has potential, how to hit upon a good pioneer who can make the campaign go viral.

In [43], we present a diffusion model developed by enriching the generalized random graph (a.k.a. configuration model) to provide insight into these questions. We offer the intuition behind the results on this model, rigorously proved in [38], and illustrate them here by taking examples of random networks having prototypical degree distributions — Poisson degree distribution, which is commonly used as a kind of benchmark, and Power Law degree distribution, which is normally used to approximate the real-world networks. On these networks, the members are assumed to have varying attitudes towards propagating the information. We analyze three cases, in particular — (1) Bernoulli transmissions, when a member influences each of its friend with probability p; (2) Node percolation, when a member influences all its friends with probability p and none with probability $1-p$; (3) Coupon-collector transmissions, when a member randomly selects one of his friends K times with replacement.

We assume that the configuration model is the closest approximation of a large online social network, when the information available about the network is very limited. The key insight offered by this study from a firm's perspective is regarding how to evaluate the effectiveness of a marketing campaign and do cost-benefit analysis by collecting relevant statistical data from the pioneers it selects. The campaign evaluation criterion is informed by the observation that if the parameters of the underlying network and the campaign effectiveness are such that the campaign can indeed reach a significant fraction of the population, then the set of good pioneers also forms a significant fraction of the population. Therefore, in such a case, the firms can even adopt the naïve strategy of repeatedly picking and targeting some number of pioneers at random from the population. With this strategy, the probability of them picking a good pioneer will increase geometrically fast with the number of tries.

6.24. Peer-to-Peer Networks

In [12], in collaboration with I. Norros (VTT, Finland) and F. Mathieu (Bell Labs), we propose a new model for peer-to-peer networking which takes the network bottlenecks into account beyond the access. This model can cope with key features of P2P networking like degree or locality constraints together with the fact that distant peers often have a smaller rate than nearby peers. Using a network model based on rate functions, we give a closed form expression of peers download performance in the system's fluid limit, as well as approximations for the other cases. Our results show the existence of realistic settings for which the average download time is a decreasing function of the load, a phenomenon that we call super-scalability.

6.25. Stability of the bipartite matching model

In [8], we consider the bipartite matching model of customers and servers introduced by Caldentey, Kaplan and Weiss (2009). Customers and servers play symmetrical roles. There are finite sets C and S of customer and server classes, respectively. Time is discrete and at each time step one customer and one server arrive in the system according to a joint probability measure μ on C$\times$ S, independently of the past. Also, at each time step, pairs of matched customers and servers, if they exist, depart from the system. Authorized em matchings are given by a fixed bipartite graph (C, S, E$\subset$ C $\times$ S). A matching policy is chosen, which decides how to match when there are several possibilities. Customers/servers that cannot be matched are stored in a buffer. The evolution of the model can be described by a discrete-time Markov chain. We study its stability under various admissible matching policies, including ML (match the longest), MS (match the shortest), FIFO (match the oldest), RANDOM (match uniformly), and PRIORITY. There exist natural necessary conditions for stability (independent of the matching policy) defining the maximal possible stability region. For some bipartite graphs, we prove that the stability region is indeed maximal for any admissible matching policy. For the ML policy, we prove that the stability region is maximal for any bipartite graph. For the MS and PRIORITY policies, we exhibit a bipartite graph with a non-maximal stability region.

6.26. Matchings on infinite graphs

Elek and Lippner (Proc. Am. Math. Soc. 138(8), 2939–2947, 2010) showed that the convergence of a sequence of bounded-degree graphs implies the existence of a limit for the proportion of vertices covered by a maximum matching. In [6], we provide a characterization of the limiting parameter via a local recursion defined directly on the limit of the graph sequence. Interestingly, the recursion may admit multiple solutions, implying non-trivial long-range dependencies between the covered vertices. We overcome this lack of correlation decay by introducing a perturbative parameter (temperature), which we let progressively go to zero. This allows us to uniquely identify the correct solution. In the important case where the graph limit is a unimodular Galton–Watson tree, the recursion simplifies into a distributional equation that can be solved explicitly, leading to a new asymptotic formula that considerably extends the well-known one by Karp and Sipser for Erdős-Rényi random graphs.

6.27. Double-hashing thresholds via local weak convergence.

A lot of interest has recently arisen in the analysis of multiple-choice "cuckoo hashing" schemes. In this context, a main performance criterion is the load threshold under which the hashing scheme is able to build a valid hashtable with high probability in the limit of large systems; various techniques have successfully been used to answer this question (differential equations, combinatorics, cavity method) for increasing levels of generality of the model. However, the hashing scheme analysed so far is quite utopic in that it requires to generate a lot of independent, fully random choices. Schemes with reduced randomness exists, such as "double hashing", which is expected to provide similar asymptotic results as the ideal scheme, yet they have been more resistant to analysis so far. In [22], we point out that the approach via the cavity method extends quite naturally to the analysis of double hashing and allows to compute the corresponding threshold. The path followed is to show that the graph induced by the double hashing scheme has the same local weak limit as the one obtained with full randomness.

6.28. Convergence of multivariate belief propagation, with applications to cuckoo hashing and load balancing

[23] is motivated by two applications, namely generalizations of cuckoo hashing, a computationally simple approach to assigning keys to objects, and load balancing in content distribution networks, where one is interested in determining the impact of content replication on performance. These two problems admit a common abstraction: in both scenarios, performance is characterized by the maximum weight of a generalization of a matching in a bipartite graph, featuring node and edge capacities. Our main result is a law of large numbers characterizing the asymptotic maximum weight matching in the limit of large bipartite random graphs, when

the graphs admit a local weak limit that is a tree. This result specializes to the two application scenarios, yielding new results in both contexts. In contrast with previous results, the key novelty is the ability to handle edge capacities with arbitrary integer values. An analysis of belief propagation algorithms (BP) with multivariate belief vectors underlies the proof. In particular, we show convergence of the corresponding BP by exploiting monotonicity of the belief vectors with respect to the so-called upshifted likelihood ratio stochastic order. This auxiliary result can be of independent interest, providing a new set of structural conditions which ensure convergence of BP.

6.29. Bypassing correlation decay for matchings with an application to XORSAT

Many combinatorial optimization problems on sparse graphs do not exhibit the correlation decay property. In such cases, the cavity method remains a sophisticated heuristic with no rigorous proof. In [24], we consider the maximum matching problem which is one of the simplest such example. We show that monotonicity properties of the problem allows us to define solutions for the cavity equations. More importantly, we are able to identify the 'right' solution of these equations and then to compute the asymptotics for the size of a maximum matching. The results for finite graphs are self-contained. We give references to recent extensions making use of the notion of local weak convergence for graphs and the theory of unimodular networks.

As an application, we consider the random XORSAT problem which according to the physics literature has a 'one-step replica symmetry breaking' (1RSB) glass phase. We derive new bounds on the satisfiability threshold valid for general graphs (and conjectured to be tight).

6.30. Sublinear-Time Algorithms for Monomer-Dimer Systems on Bounded Degree Graphs

For a graph G, let $Z(G,\lambda)$ be the partition function of the monomer-dimer system defined by $\sum_k m_k(G)\lambda^k$, where $m_k(G)$ is the number of matchings of size k in G. In [27], we consider graphs of bounded degree and develop a sublinear-time algorithm for estimating $\log Z(G,\lambda)$ at an arbitrary value $\lambda > 0$ within additive error ϵn with high probability. The query complexity of our algorithm does not depend on the size of G and is polynomial in $1/\epsilon$, and we also provide a lower bound quadratic in $1/\epsilon$ for this problem. This is the first analysis of a sublinear-time approximation algorithm for a $\#P$-complete problem. Our approach is based on the correlation decay of the Gibbs distribution associated with $Z(G,\lambda)$. We show that our algorithm approximates the probability for a vertex to be covered by a matching, sampled according to this Gibbs distribution, in a near-optimal sublinear time. We extend our results to approximate the average size and the entropy of such a matching within an additive error with high probability, where again the query complexity is polynomial in $1/\epsilon$ and the lower bound is quadratic in $1/\epsilon$. Our algorithms are simple to implement and of practical use when dealing with massive datasets. Our results extend to other systems where the correlation decay is known to hold as for the independent set problem up to the critical activity.

6.31. Reconstruction in the Labeled Stochastic Block Model

The labeled stochastic block model is a random graph model representing networks with community structure and interactions of multiple types. In its simplest form, it consists of two communities of approximately equal size, and the edges are drawn and labeled at random with probability depending on whether their two endpoints belong to the same community or not.

It has been conjectured that this model exhibits a phase transition: reconstruction (i.e. identification of a partition positively correlated with the true partition into the underlying communities) would be feasible if and only if a model parameter exceeds a threshold.

In [25], we prove one half of this conjecture, i.e., reconstruction is impossible when below the threshold. In the converse direction, we introduce a suitably weighted graph. We show that when above the threshold by a specific constant, reconstruction is achieved by (1) minimum bisection, and (2) a spectral method combined with removal of nodes of high degree.

6.32. Spectrum Bandit Optimisation

In [26], we consider the problem of allocating radio channels to links in a wireless network. Links interact through interference, modelled as a conflict graph (i.e., two interfering links cannot be simultaneously active on the same channel). We aim at identifying the channel allocation maximizing the total network throughput over a finite time horizon. Should we know the average radio conditions on each channel and on each link, an optimal allocation would be obtained by solving an Integer Linear Program (ILP). When radio conditions are unknown a priori, we look for a sequential channel allocation policy that converges to the optimal allocation while minimizing on the way the throughput loss or *regret* due to the need for exploring sub-optimal allocations. We formulate this problem as a generic linear bandit problem, and analyze it first in a stochastic setting where radio conditions are driven by a stationary stochastic process, and then in an adversarial setting where radio conditions can evolve arbitrarily. We provide, in both settings, algorithms whose regret upper bounds outperform those of existing algorithms for linear bandit problems.

6.33. Randomized Consensus with Attractive and Repulsive Links

In [29], we study convergence properties of a randomized consensus algorithm over a graph with both attractive and repulsive links. At each time instant, a node is randomly selected to interact with a random neighbor. Depending on if the link between the two nodes belongs to a given subgraph of attractive or repulsive links, the node update follows a standard attractive weighted average or a repulsive weighted average, respectively. The repulsive update has the opposite sign of the standard consensus update. In this way, it counteracts the consensus formation and can be seen as a model of link faults or malicious attacks in a communication network, or the impact of trust and antagonism in a social network. Various probabilistic convergence and divergence conditions are established. A threshold condition for the strength of the repulsive action is given for convergence in expectation: when the repulsive weight crosses this threshold value, the algorithm transits from convergence to divergence. An explicit value of the threshold is derived for classes of attractive and repulsive graphs. The results show that a single repulsive link can sometimes drastically change the behavior of the consensus algorithm. They also explicitly show how the robustness of the consensus algorithm depends on the size and other properties of the graphs.

6.34. Continuous-time Distributed Optimization of Homogenous Dynamics

This paper explores the fundamental properties of distributed minimization of a sum of functions with each function only known to one node, and a pre-specified level of node knowledge and computational capacity. We define the optimization information each node receives from its objective function, the neighboring information each node receives from its neighbors, and the computational capacity each node can take advantage of in controlling its state. It is proven that there exist a neighboring information way and a control law that guarantee global optimal consensus if and only if the solution sets of the local objective functions admit a nonempty intersection set for fixed strongly connected graphs. Then we show that for any tolerated error, we can find a control law that guarantees global optimal consensus within this error for fixed, bidirectional, and connected graphs under mild conditions. For time-varying graphs, we show that optimal consensus can always be achieved as long as the graph is uniformly jointly strongly connected and the nonempty intersection condition holds. The results illustrate that nonempty intersection for the local optimal solution sets is a critical condition for successful distributed optimization for a large class of algorithms.

6.35. Two-target Algorithms for Infinite-Armed Bandits with Bernoulli Rewards

In [16], we consider an infinite-armed bandit problem with Bernoulli rewards. The mean rewards are independent, uniformly distributed over $[0, 1]$. Rewards 0 and 1 are referred to as a success and a failure, respectively. We propose a novel algorithm where the decision to exploit any arm is based on two successive targets, namely, the total number of successes until the first failure and until the first m failures, respectively, where m is a fixed parameter. This two-target algorithm achieves a long-term average regret in $\sqrt{2n}$ for a

large parameter m and a known time horizon n. This regret is optimal and strictly less than the regret achieved by the best known algorithms, which is in $2\sqrt{n}$. The results are extended to any mean-reward distribution whose support contains 1 and to unknown time horizons. Numerical experiments show the performance of the algorithm for finite time horizons.

FORMES Team

6. New Results

6.1. Type and rewriting theory

Participants: Frédéric Blanqui, Jean-Pierre Jouannaud, Jianqi Li, Qian Wang.

Qian Wang and Bruno Barras have proved the strong normalization property of CoqMTU in presence of strong elimination, a major step towars the full certification of CoqMTU [16].

Jouannaud and Li have developped a new framework, Normal Abstract Rewriting Systems (NARS), that captures all known Church-Rosser results in presence of a termination assumption allowing to reduce equality of terms to a simpler equality on their normal forms. This result applies to the paticular case of higher-order rewriting for which it solved long-standing open problems [10].

Jouannaud and Liu have continued their investigation of Church-Rosser properties of non-terminating rewrite systems [10], showing recently first, that many results found in the litterature could be captured, and generalized, by using van Oostrom's decreasing diagram technique (accepted at Symposium on Algebraic Specifications, Kanazawa, Japan, April 2014). The next step, which has been recently completed, is a powerful result generalizing Knuth and Bendix confluence test to non terminating rewrite system (submitted).

Frédéric Blanqui, Jean-Pierre Jouannaud and Albert Rubio (Technical University of Catalonia) have developed a method aiming at carrying out termination proof for higher-order calculi. CPO appears to be the ultimate improvement of the higher-order recursive path ordering (HORPO) [25] in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard, therefore explaining the name of the improved ordering. It has been shown that CPO allows to consider higher-order rewrite rules in a simple type discipline with inductive types, that most of the guards present in the recursive calls of its core definition cannot be relaxed in any natural way without losing well-foundedness, and that the precedence on function symbols cannot be made more liberal anymore. This result is submitted to journal, and has been concurrently generalized to higher-order calculi with dependent types by Jouannaud and Li (submitted).

Frédéric Blanqui worked on the formalization in the Coq proof assistant of various definitions of the notion of α-equivalence on pure λ-terms. In particular, he formalized and formally proved equivalent the definitions of Church (1932), Curry and Feys (1958), Krivine (1993), and Gabbay and Pitts (1999). This work is freely available from the CoLoR library released on December 13th.

Frédéric Blanqui worked with John Steinberger (Tsinghua University) on the formal verification in Coq of proofs of theorems on coset arrays and non-negative integer linear combinations.

6.2. Automated theorem proving

Participant: Kim-Quyen Ly.

Kim-Quyen Ly extended her formally-proved (in Coq) automated termination-certificate (for first-order rewrite systems) verifier Rainbow for dealing with certificates using arguments filtering [22] and other termination techniques.

6.3. Simulation

Participants: Vania Joloboff, Antoine Rouquette, Shenpeng Wang.

There exists very fast Loosely Timed simulators such as **SimSoC** that can run the application software to validate its functionality and possibly test real time software using timers. But such simulators do not provide good enough timings to evaluate the software performance. The idea of "Approximately Timed" simulation is to provide a fast simulation that can be used by software developers, and yet provide performance estimate. The goal of approximately timed simulation is to provide estimates that are within a small margin error from the real hardware performance, but at a simulation speed that is an order of magnitude faster than a cycle accurate one.

Modern processors have complex architectures. They can execute a certain number of instructions per clock cycle. There are however several cases where the instruction flow is disrupted, introducing delays in the computation. In order to make an Approximately Timed simulator, our idea is to simulate enough of the processes causing the delays, not simulating the exact hardware processes of the caches and pipe line and I/Os, but using a model with wich the delays can be computed with a reasonable approximation while maintaining fast simulation. Delays may also be related to bus arbitration and interconnect access. These delays are beyond the scope of our work, but can be captured by TLM (timed) transactions. In our work, we are considering only the processor model and we rely upon TLM interface to the interconnect for peripheral access to provides us with timing delays.

We have started to investigate a new approach to provide a fast Approximately Timed ISS, that does not simulate fully the hardware, yet provides good precision estimates, and does not use stastistical methods. Our approach consists in developing a higher abstraction model of the processor (than the CA models) that still executes instructions using fast SystemC/TLM code, but in parallel maintains some architecture state to measure the delays introduces by cache misses and pipe line stalls, although the pipe line is not really simulated.

6.4. Certification of a Simulator

Participants: Vania Joloboff, Jean-François Monin, Xiaomu Shi.

We have developed a correctness proof of a part of the hardware simulator **SimSoC**. This is not only an attempt to certify a simulator, but also a new experiment on the certification of non-trivial programs written in C. We have provided a formalized representation of the ARM instruction set and addressing modes in Coq. We also constructed a Coq representation of the ARM simulator in C, using the abstract syntax defined in **CompCert**.

From these two Coq representations, we have developed Coq proofs to prove the correctness of the C code, using the operational semantics of C provided by **CompCert**.

During this work, we have also improved the technology available in Coq for performing *inversions*, a kind of proof steps which heavily occurs in this context.

All of this work has been described in Xiaomu SHI PhD thesis dissertation, presented at University of Grenoble in July 2013, and at ITP 2013 conference[15].

GALLIUM Project-Team

6. New Results

6.1. Formal verification of compilers and static analyzers

6.1.1. The CompCert formally-verified compiler

Participants: Xavier Leroy, Jacques-Henri Jourdan, Robbert Krebbers.

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 part, translating the Cminor intermediate language to PowerPC assembly and reusable for source languages other than C [5], and a front-end translating 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 Caml 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 released three versions of CompCert. Version 1.13, released in March, improves conformance with the ISO C standard by defining the semantics of comparisons involving pointers "one past" the end of an array. Such comparisons used to be undefined behaviors in earlier versions of CompCert. Robbert Krebbers formalized a reasonable interpretation of the ISO C rules concerning pointers "one past" and adapted CompCert's proofs accordingly. CompCert 1.13 also features minor performance improvements for the ARM and PowerPC back-ends, notably for parameter passing via stack locations.

Version 2.0 of CompCert, released in June, re-architects the compiler back-end around the new register allocator described in section 6.1.2 . Besides improving the performance of generated code, this new allocator made it possible to add support for 64-bit integers, that is, the `long long` and `unsigned long long` data types of ISO C99. Most arithmetic operations over 64-bit integers are expanded in-line and proved correct, but a few complex operations (division, modulus, and conversions to and from floating-point numbers) are implemented as calls into library functions.

Moreover, conformance with Application Binary Interfaces was improved, especially concerning the passing of function parameters and results of type `float` (single-precision FP numbers).

Finally, CompCert 2.0 features preliminary support for debugging information. The `-g` compiler flag causes DWARF debugging information to be generated for line numbers and call stack structure. However, no information is generated yet for C type definitions and variable declarations.

Version 2.1, released in October, addresses several shortcomings of CompCert for embedded system codes, as identified by Airbus during their experimental evaluation of CompCert. In particular, CompCert 2.1 features the `_Alignas` modifier introduced in ISO C2011, to support precise control of alignment of global variables and structure fields, and uses this modifier to implement packed structures in a more robust fashion than in earlier releases. Xavier Leroy also implemented and proved correct the optimization of integer divisions by constants introduced by Granlund and Montgomery [40].

6.1.2. Register allocation with validation a posteriori

Participant: Xavier Leroy.

Register allocation (the placement of program variables in processor registers) has a tremendous impact on the performance of compiled code. However, advanced register allocation techniques are difficult to prove correct, as they involve complex algorithms and data structures. Since the beginning of the CompCert project, we chose to avoid some of these difficult proofs by performing validation *a posteriori* for part of register allocation: the IRC graph coloring algorithm invoked during register allocation is not proved correct; instead, its results are verified at every compiler run to be a correct coloring of the given interference graph, using a simple validator proved sound in Coq.

In CompCert 2.0, we push this validation-based approach further. The whole register allocator is now subject to validation a posteriori and no longer needs to be proved correct. The validator follows the algorithm invented by Rideau and Leroy [50] and further developed by Tassarotti and Leroy. It proceeds by backward dataflow analysis of symbolic equations between program variables, registers, and stack locations.

Consequently, the new register allocator for CompCert 2.0 is much more aggressive than that of CompCert 1: it features a number of optimizations that could not be proved correct in CompCert, including live-range splitting, better handling of two-address operations and other irregularities of the x86 instruction set, an improved spilling strategy, and iterating register allocation to place temporaries introduced by spilling. Moreover, the new register allocator can handle program variables of 64-bit integer types, allocating them to pairs of 32-bit registers or stack locations. The new register allocator improves the performance of generated x86 code by up to 10% on our benchmarks.

6.1.3. Formal verification of static analyzers based on abstract interpretation

Participants: Sandrine Blazy [EPI Celtique], Vincent Laporte [EPI Celtique], Jacques-Henri Jourdan, Xavier Leroy, David Pichardie [EPI Celtique].

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 should be able to handle the same large subset of the C language as the CompCert compiler; support a combination of abstract domains, including relational domains; and 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 codes.

This year, Jacques-Henri Jourdan worked on numerical abstract domains for the static analyzer. First, he designed, programmed and proved correct an abstraction layer that transforms any relational abstract domain for mathematical, arbitrary-precision integers into a relational abstract domain for finite-precision machine integers, taking overflow and "wrap-around" behaviors into account. This domain transformer makes it possible to design numerical domains without taking into account the finiteness of machine integers. Then, he implemented and proved sound non-relational abstract domains for intervals of integers and of floating-point numbers, supporting almost all CompCert arithmetic operations.

In collaboration with team Celtique, we studied which intermediate languages of the CompCert C compiler are suitable as source language for the static analyzer. Early work by Blazy, Laporte, Maroneze and Pichardie [36] performs abstract interpretation over the RTL intermediate language, a simple language with unstructured control (control-flow graph). However, this language is too low-level to support reporting alarms at the level of the source C program.

Later this year, we decided to use the C#minor intermediate language of CompCert as source language for analysis. This language has mostly structured control (if/then/else, C loops, and `goto`), and is much closer to the source C program. Then, Jacques-Henri Jourdan, Xavier Leroy and David Pichardie designed a generic abstract interpreter for the C#minor language, parameterized by an abstract domain of execution states, using structured fixpoint iteration for loops and a function-global iteration for `goto`. Jacques-Henri Jourdan is in the process of proving the soundness of this abstract interpreter in Coq.

6.1.4. Formalization of floating-point arithmetic

Participants: Sylvie Boldo [EPI Toccata], Jacques-Henri Jourdan, Xavier Leroy, Guillaume Melquiond [EPI Toccata].

Last year, we replaced the axiomatization of floating-point numbers and arithmetic operations used in early versions of CompCert by a fully-formal Coq development, building on the Coq formalization of IEEE-754 arithmetic provided by the Flocq library of Sylvie Boldo and Guillaume Melquiond. A paper describing this work was presented at the ARITH 2013 conference [15].

This year, we extended this formalization of floating-point arithmetic with a more precise modeling of "Not a Number" special numbers, reflecting the signs and payloads of these numbers into their bit-level, in-memory representation. We also proved correct more algebraic identities over FP computations, such as $x/2^n = x \times 2^{-n}$ if $|n| < 1023$, as well as nontrivial implementation schemes for conversions between integer and FP numbers, whose correctness rely on subtle properties of the "round to odd" rounding mode. These extensions are described in a draft journal paper under submission [29], and integrated in version 2.1 of CompCert.

6.1.5. Formal verification of hardware synthesis

Participants: Thomas Braibant, Adam Chlipala [MIT].

Verification of hardware designs has been thoroughly investigated. Yet, obtaining provably correct hardware of significant complexity is usually considered challenging and time-consuming. Hardware synthesis aims to raise the level of description of circuits, reducing the effort necessary to produce them. This yields two opportunities for formal verification: a first option is to verify (part of) the hardware compiler; a second option is to study to what extent these higher-level design are amenable to formal proof.

Continuing work started during a visit at MIT under the supervision of Adam Chlipala, Thomas Braibant worked on the implementation and proof of correctness of a prototype hardware compiler. This compiler produces descriptions of circuits in RTL style from a high-level description language inspired by BlueSpec. Formal verification of hardware designs of mild complexity was conducted at the source level, making it possible to obtain fully certified RTL designs. A paper describing this compiler and two examples of certified designs was presented at the CAV 2013 conference [16].

6.2. Language design and type systems

6.2.1. The Mezzo programming language

Participants: Jonathan Protzenko, François Pottier, Thibaut Balabonski, Armaël Guéneau, Cyprien Mangin.

In the past ten years, the type systems community and the separation logic community, among others, have developed highly expressive formalisms for describing ownership policies and controlling side effects in imperative programming languages. In spite of this extensive knowledge, it remains very difficult to come up with a programming language design that is simple, effective (it actually controls side effects!) and expressive (it does not force programmers to alter the design of their data structures and algorithms).

The Mezzo programming language aims to bring new answers to these questions.

This year, we:

- made significant progress on the proof of soundness, by rewriting it in a more modular fashion;
- improved the implementation, by formalizing the algorithms and rewriting significant parts of the type-checker;
- hosted two interns who explored arithmetic reasoning and modeling of the iterator protocol, respectively;
- formalized libraries for concurrent programming in Mezzo;
- wrote both an interpreter and a compiler for the language.

A paper on Mezzo appeared in the ICFP 2013 conference [21].

During the previous year (2012), François Pottier wrote a formal definition of Mezzo, and proved that Mezzo is type-safe: that is, well-typed programs cannot crash. The proof was machine-checked using Coq. This year, Thibaut Balabonski and François Pottier extended this formalization with support for concurrency and dynamically-allocated locks, and proved that well-typed programs not only cannot crash, but also are data-race free.

The structure of the proof was re-worked so as to make it more modular. A paper, which emphasizes this modularity, has been submitted for presentation at a conference.

The new concurrent features have been integrated in the core library of Mezzo by Thibaut Balabonski. Further concurrent libraries have been included to provide more communication primitives, such as channels for message passing.

Jonathan Protzenko worked on formalizing the type-checking algorithms currently used in the Mezzo prototype compiler. This led to practical results in the form of improvements to the type-checker: we now type-check more programs, and the success of the type-checker is more predictable as well. Some soundness bugs have been identified and fixed. The design of some of the language's features has been improved as well.

The formalization of the type-checker was presented at the IFL 2013 conference, and is to appear in the post-symposium proceedings in 2014.

We set out to promote Mezzo in the wild. Protzenko packaged the software to make it available widely via OPAM, wrote a tutorial for end-users [34], communicated through blog posts about the language, and released the source code online for others to contribute.

We also spread the word about Mezzo through various seminar talks and discussions with other teams (Carnegie-Mellon university, Cambridge Computer Lab, Aarhus University, Brasilia University), and by communicating in international conferences (ICFP'13, FSFMA'13).

This year, two interns worked with us on Mezzo. Armaël Guéneau (L3; June-July 2013) and Cyprien Mangin (M1; April-July 2013) explored several experimental aspects of the language. In particular, Armaël worked on an encoding of iterators in an object-oriented style, which involves transfers of ownership and typestate changes; while Cyprien improved the treatment of arrays and implemented an experimental extension of Mezzo with arithmetic assertions. Armaël presented his work at the workshop HOPE 2013. This work is also described in a short unpublished paper [33].

6.2.2. System F with coercion constraints

Participants: Julien Cretin, Didier Rémy.

Expressive type systems often allow non trivial conversions between types, which may lead to complex, challenging, and sometimes ad hoc type systems. Such examples are the extension of System F with type equalities to model GADTs and type families of Haskell, or the extension of System F with explicit contracts. A useful technique to simplify the meta-theoretical studies of such systems is to make type conversions explicit as "coercions" inside terms.

Following a general approach to coercions, we extended System F with a richer type-level language and a proposition language. Propositions contain a first-order logic, a coinduction mechanism, coherence assertions and coercion assertions. Types are classified by kinds and extended in order to handle lists of types. We introduce a particular kind restricting a previous kind to its types satisfying a proposition. Abstracting over such a kind means abstracting over arbitrary propositions, and thus enables coercion abstraction. Type abstraction must be coherent: the kind of the abstract type has to be inhabited by a witness type. This language, called Fcc, extends our previous language parametric F-iota and additionally subsumes Constraint ML.

We also extended Fcc with incoherent polymorphism in order to model GADTs. Unlike coercions and thus coherent polymorphism, incoherent polymorphism is not erasable. But in counterpart, incoherent abstraction does not require the kind to be inhabited. Since abstracting over incoherent types permits to write unsound terms, incoherent abstraction has to block the reduction of terms.

This work is part of Julien Cretin's Ph.D. dissertation [11], which will be defended in January 2014.

6.2.3. Type inference for GADTs

Participants: Jacques Garrigue [Nagoya University], Didier Rémy.

Type inference for generalized algebraic data types (GADTs) is inherently non monotone: assuming more specific types for GADTs may ensure more invariants, which may result in more general types. This is problematic for type inference and some amount of type annotations is required.

Moreover, even when types of GADTs parameters are explicitly given, they introduce equalities between types, which makes them inter-convertible but with a limited scope. This may create an ambiguity when leaving the scope of the equation: which element should be used for representing the equivalent forms? Idealy, one should use a type disjunction, but this is not allowed—for good reasons. Hence, to avoid arbitrary choices, these situations must be rejected as ambiguous, forcing the user to write more annotations to resolve the ambiguities.

We proposed a new approach to type inference with GADTs. While some uses of equations are unavoidable and create *real* ambiguities, others are gratuitous and create *artificial* ambiguities, To distinguish between the two we introduced *ambivalent types*, which are a way to trace unavoidable uses of equations within types themselves. We then redefined ambiguities so that only ambivalent types become ambiguous and should be rejected or resolved by a programmer annotation. Interestingly, this solution is fully compatible with unification-based type inference algorithms used in ML dialects.

This work was presented at the APLAS 2013 conference [20]. It is also implemented in the OCaml language since version 4.00.

6.2.4. GADTs and Subtyping

Participants: Gabriel Scherer, Didier Rémy.

Following the addition of GADTs to the OCaml language in version 4.00 released this year, we studied the theoretical underpinnings of variance subtyping for GADTs. The question is to decide which variances should be accepted for a GADT-style type declaration that includes type equality constraints in constructor types. This question exposes a new notion of decomposability and unexpected tensions in the design of a subtyping relation. A paper describing our formalization was presented at the ESOP 2013 conference [23].

6.2.5. Singleton types for code inference

Participants: Gabriel Scherer, Didier Rémy.

We continued working on the use of singleton types for code inference. If we can prove that a type contains, in a suitably restricted pure lambda-calculus, a unique inhabitant modulo program equivalence, the compiler can infer the code of this inhabitant. This opens the way to type-directed description of boilerplate code, through type inference of finer-grained type annotations. As this is still work in progress, there was no publication on this topic this year, but we presented our directions on three occasions: at the PLUME team in ENS Lyon, at the LIX team in École Polytechnique (whose proof-search research is highly relevant to our work), and at the Dependently Typed Programming workshop (satellite of the International Conference on Functional Programming) in Boston.

6.2.6. Open closure types

Participants: Gabriel Scherer, Jan Hoffmann [Yale University, FLINT group].

During a visit to Yale, Gabriel Scherer worked with Jan Hoffmann on a type system for program analysis of higher-order functional languages. Open closure types are a novel typing construct that lets the type system statically reason about closure variables present in the lexical context. This allows fine-grained analysis (e.g., for resource consumption or information-flow control) of functional programming patterns such as function currying. This work was presented at the LPAR 2013 conference [22] (Logic for Programming, Artificial Intelligence, and Reasoning) in October.

6.3. Shared-memory parallelism

6.3.1. Algorithms and data structures for parallel computing

Participants: Umut Acar, Arthur Charguéraud [EPI Toccata], Mike Rainey.

The ERC Deepsea project, with principal investigator Umut Acar, started in June and is hosted by the Gallium team. This project aims at developing techniques for parallel and self-adjusting computations 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 in the previous three years. 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 have recently been pursuing two main lines of work.

We have been developing an algorithm that is able to perform dynamic load balancing in the style of work stealing but without requiring atomic read-modify-write operations. These operations may scale poorly with the number of cores due to synchronization bottlenecks. We have designed the algorithm, proved it correct using a new technique for the x86-TSO weak memory model. We have evaluated our algorithm on a modern multicore machine. Although we use no synchronization operations, we achieve performance that is no more than a few percent slower than the industrial-strengh algorithm, even though the industrial-strength algorithm takes full advantage of synchronization operations. We have a soon-to-be-submitted research article describing our contributions [25].

The design of efficient parallel graph algorithms requires a sequence data structure that supports logarithmic-time split and concatenation operations in addition to push and pop operations with excellent constant factors. We have designed such a data structure by building on a recently introduced data structure called Finger Tree and by integrating a "chunking" technique. Our chunking technique is based on instantiating the leaves of the Finger Tree with chunks of contiguous memory. Unlike previous chunked data structures, we are able to prove efficient constant factors even in worst-case scenarios. Moreover, we implemented our data structure in C++ and OCaml and showed it to be competitive with state-of-the-art sequence data structures that do not support split and concatenation operations. We are currently writing a report on our results.

6.3.2. Weak memory models

Participants: Luc Maranget, Jacques-Pascal Deplaix, Jade Alglave [University College London].

Modern multicore and multiprocessor computers do not follow the intuitive "Sequential Consistency" model that would define a concurrent execution as the interleaving of the execution of its constituting 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 instruction and to the presence of sophisticated (and cooperating) caching devices between processors and memory.

In the last few years, Luc Maranget took part in an international research effort to define the semantics of the computers of the multi-core era. This research effort relies both on formal methods for defining the models and on intensive experiments for validating the models. Joint work with, amongst others, Jade Alglave (now at University College London) and Peter Sewell (University of Cambridge) achieved several significant results, including two semantics for the IBM Power and ARM memory models: one of the operational kind [52] and the other of the axiomatic kind [46]. In particular, Luc Maranget is the main developer of the **diy** tool suite (see section 5.3). Luc Maranget also performs most of the experiments involved.

In 2013, Luc Maranget pursued this collaboration. He mainly worked with Jade Alglave to produce a new model for Power/ARM. The new model is simpler than the previous ones, in the sense that it is based on fewer mathematical objects and can be simulated more efficiently than the previous models. The new model is at the core of a journal submission which is now at the second stage of reviewing. The submitted work contains in-depth testing of ARM devices which led to the discovery of anomalous behaviours acknowledged as such by our ARM contact, and of legitimate features now included in the model. The new model also impacted our **diy** tool suite that now includes a generic memory model simulator built by following the principles exposed in the submitted article. At the moment the new simulator is available as an experimental release (http://diy.inria.fr/herd). It will be include in future releases of the tool suite.

In the same research theme, Luc Maranget supervises the internship of Jacques-Pascal Deplaix (EPITECH), from Oct. 2013 to May 2014. The internship aims at extending **litmus**, our tool to to run tests on hardware: at the moment **litmus** accepts test written in assembler; Jacques-Pascal is extending **litmus** so that it accepts tests written in C. The general objective is to achieve conformance testing of C compilers and machines with respect to the new C11/C++11 standard.

6.4. The OCaml language and system

6.4.1. The OCaml system

Participants: Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant, Xavier Leroy, Gabriel Scherer.

This year, we released version 4.01.0 of the OCaml system. This is a major release that fixes about 140 bugs and introduces 44 new features suggested by users. Damien Doligez acted as release manager for this version.

The major innovations in OCaml 4.01 are:

- The overloading of variant constructors and record field labels, resolved using typing information. Before this, programmers had to use globally unique field labels across all record types. The new typechecking algorithm enables programmers to use more natural names for fields in their data structures. The algorithm is carefully engineered to preserve principality of inferred types.
- New warnings give the programmer the option of applying very strict checking of problematic constructs in the source code.

Other features of this release include:

- Suggestion of possible typos in case of "unbound identifier" error.
- New infix application operators in the standard library.
- Options to reduce the verbosity (and enhance the readability) of error messages.
- Many internal improvements, especially in compiler performance.

In parallel, we designed and experimented with several new features that are candidate for inclusion in the next major release of OCaml in 2014:

- Module aliases: a more efficient way of typechecking and compiling module declarations of the form `module M = ModuleName`, providing a lighter, more practical alternative to packed modules and reducing the need for name spaces.
- Extension points and preprocessing by rewriting abstract syntax trees: this approach provides an alternative to Camlp4 for macro processing and automatic code generation.
- A native code generator for the new ARM 64 bit instruction set (also known as AArch64).
- Several ongoing experiments to improve the performance of OCaml-compiled code: more aggressive function inlining and constant propagation; more unboxing of numbers; and a pass of common subexpression elimination.

6.4.2. Run-time types for the OCaml language

Participants: Grégoire Henry, Jacques Garrigue [University of Nagoya], Fabrice Le Fessant.

With the addition of GADTs to OCaml in version 4.00, it is now possible to provide a clean implementation of run-time types in the language, thus allowing the definition of polytypic function, a.k.a. generic function defined by case analysis on the structure of its argument's type. However, when integrating this mechanism into the language, its interaction with other parts of the type-system proved delicate, the main difficulty being the semantic of abstract types.

In collaboration with Jacques Garrigue during a 3 month stay in Japan, Grégoire Henry worked on different semantics for the runtime representation of abstract types. They tried to design a mechanism that preserves abstraction by default, and still allows to propagate type information when requested by the programmer.

6.4.3. Multi-runtime OCaml

Participants: Luca Saiu, Fabrice Le Fessant.

Multicore architectures are now broadly available, and developers expect their programs to be able to benefit from them. In OCaml, there is no portable way to use such architectures, as only one OCaml thread can run at any time.

As part of the ANR project "BWare", Luca Saiu and Fabrice Le Fessant developed a multi-runtime version of OCaml that takes advantage of multicore architectures. In this version, a program can start several runtimes that can run on different cores. As a consequence, OCaml threads running on different runtimes can run concurrently. This implementation required a lot of rewriting of the OCaml runtime system (written in C), to make all global variables context-dependent and all functions reentrant. The compiler was also modified to generate reentrant code and context-dependent variables. The sources of the prototype were released in September 2013, to be tested by users.

Luca Saiu then developed a library based on skeletons to facilitate the development of parallel applications that take advantage of the multi-runtime architecture.

6.4.4. Evaluation strategies and standardization

Participants: Thibaut Balabonski, Flávio de Moura [Universidade de Brasília].

During the past years, Thibaut Balabonski studied evaluation strategies, laziness and optimality for functional programming languages, in particular in relation to pattern matching. These investigations continued this year, with two highlights:

- Publication in the ICFP conference [14] of a theoretical result relating fully lazy evaluation (as can be found in some Haskell compilers) to optimal reduction in the weak λ-calculus.
- Collaboration with Flávio de Moura (Universidade de Brasília) on so-called "standard" evaluation strategies for a calculus with rich pattern matching mechanisms (the *Pure Pattern Calculus* of Jay and Kesner [42]). The challenge here lies in that the calculus does not satisfies the usual stability properties. As a consequences, standard strategies are not unique anymore, and new approaches are needed. A paper is in preparation.

6.5. Software specification and verification

6.5.1. Tools for TLA+

Participants: Damien Doligez, Jael Kriener, Leslie Lamport [Microsoft Research], Stephan Merz [EPI VeriDis], Tomer Libal [Microsoft Research-Inria Joint Centre], Hernán Vanzetto [Microsoft Research-Inria Joint Centre].

Damien Doligez is head of the "Tools for Proofs" team in the Microsoft-Inria Joint Centre. The aim of this team is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing the ideas in [43], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, the TLA+ tools were released as open-source (MIT license), and in September we released a new version of the TLA+ Proof System (TLAPS), an environment for writing and checking TLA+ proofs. This environment is described in [38].

We have implemented a (not yet released) extension of TLAPS to deal with proofs of temporal formulas, using the propositional temporal logic prover LS4 as a back-end. Until now, TLAPS could only be used to prove safety properties (invariants). With this new version, our users will be able to prove liveness properties (absence of deadlock), refinement relations between specifications, etc.

Jael Kriener started a 2-year post-doc contract in December. She is working on theoretical and implementation aspects of TLA+ and TLAPS.

Web sites:
http://research.microsoft.com/users/lamport/tla/tla.html
http://tla.msr-inria.inria.fr/tlaps

6.5.2. The Zenon automatic theorem prover

Participants: Damien Doligez, David Delahaye [CNAM], Pierre Halmagrand [CNAM], Olivier Hermant [Mines ParisTech], Mélanie Jacquel [CNAM].

Damien Doligez continued the development of Zenon, a tableau-based prover for first-order logic with equality and theory-specific extensions.

David Delahaye and Mélanie Jacquel designed and implemented (with some help from Damien Doligez) an extension of Zenon called SuperZenon, based on the Superdeduction framework of Brauner, Houtmann, and Kirchner [37]. Mélanie Jacquel defended her thesis on this subject in April.

Pierre Halmagrand did an internship and started a thesis on integrating Deduction Modulo in Zenon; some results of this work are described in two papers published at LPAR [19] and IWIL [18].

6.5.3. Implementing hash-consed structures in Coq

Participants: Thomas Braibant, Jacques-Henri Jourdan, David Monniaux [CNRS, VERIMAG].

Hash-consing is a programming technique used to implement maximal sharing of immutable values in memory, keeping a single copy of semantically equivalent objects. Hash-consed data-structures give a unique identifier to each object, allowing fast hashing and comparisons of objects. This may lead to major improvements in execution time by itself, but it also make it possible to do efficient memoization of computations.

Hash-consing and memoization are examples of imperative techniques that are of prime importance for performance, but are not easy to implement and prove correct using the purely functional language of a proof assistant such as Coq. In a joint article at ITP 2013 [17], we described three different implementation techniques for hash-consed data-structures in Coq through the running example of Binary Decision Diagrams (BDDs). BDDs are representations of Boolean functions, and are often used in software and hardware verification tools (e.g., model checkers).

We substantially improved the work described in this ITP 2013 article afterwards. First, we came up with a fourth implementation technique for hash-consed data-structures in Coq. Then, we performed an in-depth comparative study of how our "design patterns" for certified hash-consing fare on two real-scale examples: BDDs and lambda-terms. This work is currently under revision for publication in a journal.

6.5.4. Working with names and binders

Participant: François Pottier.

François Pottier released **dblib**, a Coq library that helps work with de Bruijn indices in a generic and lightweight manner. This library is used in the formalization of Mezzo (see section 6.2.1). It is available at http://gallium.inria.fr/~fpottier/.

6.6. Technology transfer

6.6.1. Analysis of the Scilab Language

Participants: Fabrice Le Fessant, Michael Laporte.

The Scilab language is a scripting language providing easy access to efficient implementations of mathematical operations (on matrices, for example). It suffers from the lack of verifications of an untyped language, together with the performance problems of an interpreted language. As part of the FUI Richelieu project, Fabrice Le Fessant and Michael Laporte have been investigating solutions to these issues.

The first part of the work was to clarify the semantics of the Scilab language. For that, an interpreter was implemented in OCaml, based on the C++ AST provided by the forthcoming version 6 of Scilab. This work exhibited a number of bugs in the new implementation, and proved to be more performant than the C++ implementation, thanks to a better algorithm to manage the dynamic scopes of Scilab.

The second part of the work was to understand how users write Scilab code. For that, a style-checking application, called *Scilint*, has been developed. It implements static checking of some properties of Scilab programs, to be able to detect runtime errors before running the program. Warnings are displayed for suspicious cases. Using Scilint on large sets of Scilab code (from the Scilab forge or the Atom repository) showed that the most erroneous features of Scilab are commonly used and that, to achieve the ultimate goal of partial typing of the language, a subset of the language must be specified that the user should conform to, in order for the code to benefit from the next part of the work, i.e. just-in-time compilation.

GAMMA3 Project-Team

4. New Results

4.1. From discrete to continuous metric fields

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Adaptive computation using adaptive meshes is now recognized as essential for solving complex PDE problems. This computation requires at each step the definition of a continuous metric field to govern the generation of the adapted meshes. In practice, via an appropriate *a posteriori* error estimation, metrics are calculated at the vertices of the computational domain mesh. In order to obtain a continuous metric field, the discrete field is interpolated in the whole domain mesh. In this study, a new method for interpolating discrete metric fields, based on a so-called "natural decomposition" of metrics, is introduced. The proposed method is based on known matrix decompositions and is computationally robust and efficient. Some qualitative comparisons with classical methods are made to show the relevance of this methodology [19].

4.2. Hex-dominant meshing of geologic structures

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Simulation by a finite volume method of the transfer by water of radioactive elements in sites of nuclear waste storage, on large time and space scales, is the only possible way to analyze the safety of disposal. To properly represent the different pathways of radionuclides, surface topography (valleys, reliefs, rivers), geologic layers and simplified storage facilities must be accurately modeled. We propose a new methodology for generating hex-dominant meshes (well suited for a finite volume formulation) of geologic structures complying with these different geometric constraints.

First, a reference 2D domain is obtained by projecting all the line constraints into a horizontal plane. Different size specifications are given for workings, outcrop lines and rivers. Using an adaptive methodology, the size variation is bounded by a specified threshold in order to obtain a high quality quad-dominant mesh. Secondly, a hex-dominant mesh of the geological medium is generated by a vertical extrusion. Depending on the configuration of the surfaces found (interfaces between two layers, top or bottom faces of underground workings), hexahedra, prisms, pyramids and tetrahedra are generated. The generation of volume elements follows a global order established on the whole set of surfaces to ensure the conformity of the resulting mesh. An example of mesh construction of a geologic structure illustrates the suitability of the proposed methodology [22].

4.3. Applications du maillage et développements de méthodes avancées pour la cryptographie

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Michael François
Validité du projet: 2009-2013.
Production scientifique: 1 thèse soutenue (M. François, 17/10/2012), 6 articles publiés.
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 développé.

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.

4.4. Développement de méthodes avancées et maillages appliqués à l'étude de la nanomorphologie des nanotubes/fils en suspension liquide"

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Abel Cherouat, Houman Borouchaki, Laurence Giraud-Moreau, Anis Chaari.
Validité du projet: 2011-2014.
Production scientifique: 1 thèse en cours (A. Chaari), 1 articles publiés, 1 conférence (CSMA 2013).
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).

4.5. Applications du maillage à des problèmes multi-physiques, développement de méthodes de résolutions avancées et modélisation électromagnetisme-thermique-mécanique à l'échelle mesoscopique

Participants: Dominique Barchiesi [correspondant], Abel Cherouat, Thomas Grosges, Houman Borouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari, Fadhil Mezghani
Validité du projet: 2009-2015 (thèse de Fadhil Mezghani initiée en 2012 coencadrée par D. Barchiesi et A. Cherouat).
Production scientifique: 1 thèse soutenue (S. Kessentini, 22/10/2012), 9 articles publiés, 4 conférences.

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 2013 a constitué en la poursuite de la pré-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.

4.6. Validity of rational and nonrational Lagrange finite elements of degree 1 and 2

Participants: Paul-Louis George [correspondant], Houman Borouchaki.

A finite element is valid if its jacobian is strictly positive everywhere. The jacobian is the determinant of the jacobian matrix related to the partials of the mapping function which maps the parameter space (reference element) to the current element. Apart when it is constant, the jacobian is a polynomial whose degree is related to the degree of the finite element (but not the same in general). The value of the jacobian varies after the point where it is evaluated. Validating an element relies in finding the sign of this polynomial when one traverses the element.

Various papesr and a synthesis of those reports, shows how to calculating the jacobian of the different usual Lagrange finite elements of degree 1 and 2. To this end, we take the form of this polynomial as obtained in the classical finite element framework (shape functions and nodes) or after reformulating the element by means of a Bezier form (Bernstein polynomials and control points) which makes easier the discussion. We exhibit sufficient (necessary and sufficient in some cases) conditions to ensure the validity of a given element.

4.7. Mesh adaptation for very high-order numerical scheme

Participants: Frédéric Alauzet [correspondant], Adrien Loseille, Estelle Mbinky.

In the past, we have demonstrated that multi-scale anisotropic mesh adaptation is a powerful tool to accurately simulate compressible flow problems and to obtain faster convergence to continuous solutions. But, this was limited to second order numerical scheme. Nowadays, numerous teams are working on the development of very high-order numerical scheme (e.g. of third or greater order): Discontinous Galerkin, Residual Distribution scheme, Spectral method, ...

This work extend interpolation error estimates to higher order numerical solution representation. We have examined the case of third-order accuracy. The first step is to reduce the tri-linear form given by the third order error term into a quadratic form based on the third order derivative. From this local error model, the optimal mesh is exhibited thanks to the continuous mesh framework.

4.8. Visualization and modification of high-order curved meshes

Participants: Julien Castelneau, Adrien Loseille [correspondant], Loïc Maréchal.

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 has been addressed. The software VIZIR is freely available at https://www.rocq.inria.fr/gamma/gamma/vizir/.

4.9. A changing-topology ALE numerical scheme

Participants: Frédéric Alauzet [correspondant], Nicolas Barral.

The main difficulty arising in numerical simulations with moving geometries is to handle the displacement of the domain boundaries, *i.e.,* the moving bodies. Only vertices displacement is not sufficient to achieve complex movement such as shear. We proved that the use of edge swapping allows us to achieve such complex displacement. We therefore developed an ALE formulation of this topological mesh modification to preserve the solver accuracy and convergence order. The goal is to extend to 3D the previous work done in 2D.

4.10. 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.

4.11. Serial and parallel cavity-based mesh adaptation

Participants: 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 along with adaptive quasi-structured elements. The later point is of main interest when dealing with viscous phenomena where a boundary layer mesh is needed [26].

In addition, a parallel version of the algorithm was designed [27].

GANG Project-Team

5. New Results

5.1. Understanding graph representations

5.1.1. Connected graph searching

5.1.1.1. Computing H-Joins with Application to 2-Modular Decomposition

Participants: Michel Habib, Antoine Mamcarz, Fabien de Montgolfier.

We present in [10], a general framework to design algorithms that compute H-join. For a given bipartite graph H, we say that a graph G admits a H-join decomposition or simply a H-join, if the vertices of G can be partitioned in |H| parts connected as in H. This graph H is a kind of pattern, that we want to discover in G. This framework allows us to present fastest known algorithms for the computation of P 4-join (aka N-join), P 5-join (aka W-join), C 6-join (aka 6-join). We also generalize this method to find a homogeneous pair (also known as 2-module), a pair M 1,M 2 such that for every vertex x$\notin$(M 1$\cup$M 2) and i$\in$1,2, x is either adjacent to all vertices in M i or to none of them. First used in the context of perfect graphs (Chvátal and Sbihi in Graphs Comb. 3:127-139, 1987), it is a generalization of splits (a.k.a. 1-joins) and of modules. The algorithmics to compute them appears quite involved. In this paper, we describe an O(mn 2)-time algorithm computing all maximal homogeneous pairs of a graph, which not only improves a previous bound of O(mn 3) for finding only one pair (Everett et al. in Discrete Appl. Math. 72:209-218, 1997), but also uses a nice structural property of homogenous pairs, allowing to compute a canonical decomposition tree for sesquiprime graphs (i.e., graphs G having no module and such that for every vertex v$\in$G, G-v also has no module).

5.1.1.2. Algorithmic Aspects of Switch Cographs

Participants: Vincent Cohen-Addad, Michel Habib, Fabien de Montgolfier.

The paper [27], introduces the notion of involution module, the first generalization of the modular decomposition of 2-structure which has a unique linear-sized decomposition tree. We derive an $O(n^2)$ decomposition algorithm and we take advantage of the involution modular decomposition tree to state several algorithmic results. Cographs are the graphs that are totally decomposable w.r.t modular decomposition. In a similar way, we introduce the class of switch cographs, the class of graphs that are totally decomposable w.r.t involution modular decomposition. This class generalizes the class of cographs and is exactly the class of (Bull, Gem, Co-Gem, C_5)-free graphs. We use our new decomposition tool to design three practical algorithms for the maximum cut, vertex cover and vertex separator problems. The complexity of these problems was still unknown for this class of graphs. This paper also improves the complexity of the maximum clique, the maximum independant set, the chromatic number and the maximum clique cover problems by giving efficient algorithms, thanks to the decomposition tree. Eventually, we show that this class of graphs has Clique-Width at most 4 and that a Clique-Width expression can be computed in linear time.

5.1.1.3. LDFS-Based Certifying Algorithm for the Minimum Path Cover Problem on Cocomparability Graphs

Participants: Derek Corneil, Dalton Barnaby, Michel Habib.

For graph $G(V, E)$, a minimum path cover (MPC) is a minimum cardinality set of vertex disjoint paths that cover V (i.e., every vertex of G is in exactly one path in the cover). This problem is a natural generalization of the Hamiltonian path problem. Cocomparability graphs (the complements of graphs that have an acyclic transitive orientation of their edge sets) are a well studied subfamily of perfect graphs that includes many popular families of graphs such as interval, permutation, and cographs. Furthermore, for every cocomparability graph G and acyclic transitive orientation of the edges of $\overline{G}$ there is a corresponding poset P_G; it is easy to see that an MPC of G is a linear extension of P_G that minimizes the bump number of P_G. Although there are directly graph-theoretical MPC algorithms (i.e., algorithms that do not rely on poset formulations) for various subfamilies of cocomparability graphs, notably interval graphs, until now all MPC algorithms for

cocomparability graphs themselves have been based on the bump number algorithms for posets. In this paper [5], we present the first directly graph-theoretical MPC algorithm for cocomparability graphs; this algorithm is based on two consecutive graph searches followed by a certifying algorithm. Surprisingly, except for a lexicographic depth first search (LDFS) preprocessing step, this algorithm is identical to the corresponding algorithm for interval graphs. The running time of the algorithm is $O(\min(n^2, n + \text{mloglogn}))$, with the nonlinearity coming from LDFS.

5.1.1.4. Easy identification of generalized common and conserved nested intervals

Participants: Fabien de Montgolfier, Mathieu Raffinot, Irena Rusu.

In the paper [28], we explain how to easily compute gene clusters, formalized by classical or generalized nested common or conserved intervals, between a set of K genomes represented as K permutations. A $b-$nested common (resp. conserved) interval I of size $|I|$ is either an interval of size 1 or a common (resp. conserved) interval that contains another b-nested common (resp. conserved) interval of size at least $|I| - b$. When $b = 1$, this corresponds to the classical notion of nested interval. We exhibit two simple algorithms to output all b-nested common or conserved intervals between K permutations in $O(Kn + \text{nocc})$ time, where nocc is the total number of such intervals. We also explain how to count all b-nested intervals in O(Kn) time. New properties of the family of conserved intervals are proposed to do so.

5.1.1.5. On computing the diameter of real-world undirected graphs

Participants: Pierluigi Crescenzi, Roberto Grossi, Michel Habib, Leonardo Lanzi, Andrea Marino.

We propose in [2],a new algorithm for the classical problem of computing the diameter of undirected unweighted graphs, namely, the maximum distance among all the pairs of nodes, where the distance of a pair of nodes is the number of edges contained in the shortest path connecting these two nodes. Although its worst-case complexity is O(nm) time, where n is the number of nodes and m is the number of edges of the graph, we experimentally show that our algorithm works in O(m) time in practice, requiring few breadth-first searches to complete its task on almost 200 real-world graphs.

5.1.1.6. Toward more localized local algorithms: removing assumptions concerning global knowledge

Participants: Amos Korman, Jean-Sébastien Sereni, Laurent Viennot.

Numerous sophisticated local algorithm were suggested in the literature for various fundamental problems. Notable examples are the MIS and $(\Delta + 1)$-coloring algorithms by Barenboim and Elkin, by Kuhn, and by Panconesi and Srinivasan, as well as the $o(\Delta^2)$-coloring algorithm by Linial. Unfortunately, most known local algorithms (including, in particular, the aforementioned algorithms) are *non-uniform*, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the maximum degree Δ or the number of nodes n. This paper [11], provides a rather general method for transforming a non-uniform local algorithm into a *uniform* one. Furthermore, the resulting algorithm enjoys the same asymptotic running time as the original non-uniform algorithm. Our method applies to a wide family of both deterministic and randomized algorithms. Specifically, it applies to almost all of the state of the art non-uniform algorithms regarding MIS and Maximal Matching, as well as to many results concerning the coloring problem. (In particular, it applies to all aforementioned algorithms.) To obtain our transformations we introduce a new distributed tool called *pruning* algorithms, which we believe may be of independent interest.

5.1.2. Self-organizing Flows in Social Networks

Participants: Nidhi Hegde, Laurent Massoulié, Laurent Viennot.

Social networks offer users new means of accessing information, essentially relying on ”social filtering”, i.e. propagation and filtering of information by social contacts. The sheer amount of data flowing in these networks, combined with the limited budget of attention of each user, makes it difficult to ensure that social filtering brings relevant content to the interested users. Our motivation in this paper [24], is to measure to what extent self-organization of the social network results in efficient social filtering. To this end we introduce flow games, a simple abstraction that models network formation under selfish user dynamics, featuring user-specific interests and budget of attention. In the context of homogeneous user interests, we show that selfish dynamics converge to a stable network structure (namely a pure Nash equilibrium) with close-to-optimal information

dissemination. We show in contrast, for the more realistic case of heterogeneous interests, that convergence, if it occurs, may lead to information dissemination that can be arbitrarily inefficient, as captured by an unbounded "price of anarchy". Nevertheless the situation differs when users' interests exhibit a particular structure, captured by a metric space with low doubling dimension. In that case, natural autonomous dynamics converge to a stable configuration. Moreover, users obtain all the information of interest to them in the corresponding dissemination, provided their budget of attention is logarithmic in the size of their interest set.

5.2. Large Scale Networks Performance and Modeling

5.2.1. Can P2P Networks be Super-Scalable?

Participants: François Baccelli, Fabien Mathieu, Ilkka Norros, Rémi Varloot.

We propose in [14], a new model for peer-to-peer networking which takes the network bottlenecks into account beyond the access. This model can cope with key features of P2P networking like degree or locality constraints together with the fact that distant peers often have a smaller rate than nearby peers. Using a network model based on rate functions, we give a closed form expression of peers download performance in the system's fluid limit, as well as approximations for the other cases. Our results show the existence of realistic settings for which the average download time is a decreasing function of the load, a phenomenon that we call super-scalability.

5.2.2. Contenu généré par les utilisateurs : une étude sur DailyMotion

Participants: Yannick Carlinet, The Dang Huynh, Bruno Kauffmann, Fabien Mathieu, Ludovic Noirie, Sébastien Tixeuil.

Actuellement, une large part du trafic Internet vient de sites de "User-Generated Content" (UGC). Comprendre les caractéristiques de ce trafic est important pour les opérateurs (dimensionnement réseau), les fournisseurs (garantie de la qualité de service) et les équipementiers (conception d'équipements adaptés). Dans ce contexte, nous proposons [15], d'analyser et de modéliser des traces d'usage du site DailyMotion.

5.2.3. Rumor Spreading in Random Evolving Graphs

Participants: Andrea Clementi, Pierluigi Crescenzi, Carola Doerr, Pierre Fraigniaud, Isopi Marco, Alessandro Panconesi, Pasquale Francesco, Silvestri Riccardo.

In [13], we aim at analyzing the classical information spreading "push" protocol in *dynamic* networks. We consider the *edge-Markovian* evolving graph model which captures natural temporal dependencies between the structure of the network at time t, and the one at time $t+1$. Precisely, a non-edge appears with probability p, while an existing edge dies with probability q. In order to fit with real-world traces, we mostly concentrate our study on the case where $p = \Omega(\frac{1}{n})$ and q is constant. We prove that, in this realistic scenario, the "push" protocol does perform well, completing information spreading in $O(\log n)$ time steps, w.h.p., even when the network is, w.h.p., disconnected at every time step (e.g., when $p \ll \frac{\log n}{n}$). The bound is tight. We also address other ranges of parameters p and q (e.g., $p+q=1$ with arbitrary p and q, and $p = \Theta\left(\frac{1}{n}\right)$ with arbitrary q). Although they do not precisely fit with the measures performed on real-world traces, they can be of independent interest for other settings. The results in these cases confirm the positive impact of dynamism.

5.3. Complexity issues in distributed graph algorithms

5.3.1. What can be decided locally without identifiers?

Participants: Pierre Fraigniaud, Mika Göös, Amos Korman, Jukka Suomela.

Do unique node identifiers help in deciding whether a network G has a prescribed property P? We study this question in the context of distributed local decision, where the objective is to decide whether $G \in P$ by having each node run a constant-time distributed decision algorithm. If $G \in P$, all the nodes should output yes; if $G \notin P$, at least one node should output no. A recent work (Fraigniaud et al., OPODIS 2012) studied the role of identifiers in local decision and gave several conditions under which identifiers are not needed. In this article [21], we answer their original question. More than that, we do so under all combinations of the following two critical variations on the underlying model of distributed computing: (B): the size of the identifiers is bounded by a function of the size of the input network; as opposed to ($\neg B$): the identifiers are unbounded. (C): the nodes run a computable algorithm; as opposed to ($\neg C$): the nodes can compute any, possibly uncomputable function. While it is easy to see that under ($\neg B, \neg C$) identifiers are not needed, we show that under all other combinations there are properties that can be decided locally if and only if identifiers are present. Our constructions use ideas from classical computability theory.

5.3.2. Local Distributed Decision

Participants: Pierre Fraigniaud, Amos Korman, David Peleg.

A central theme in distributed network algorithms concerns understanding and coping with the issue of locality. Inspired by sequential complexity theory, we focus on a complexity theory for distributed decision problems. In the context of locality, solving a decision problem requires the processors to independently inspect their local neighborhoods and then collectively decide whether a given global input instance belongs to some specified language. Our paper [7], introduces several classes of distributed decision problems, proves separation among them and presents some complete problems. More specifically, we consider the standard LOCAL model of computation and define LD (for local decision) as the class of decision problems that can be solved in constant number of communication rounds. We first study the intriguing question of whether randomization helps in local distributed computing, and to what extent. Specifically, we define the corresponding randomized class BPLD, and ask whether LD=BPLD. We provide a partial answer to this question by showing that in many cases, randomization does not help for deciding hereditary languages. In addition, we define the notion of local many-one reductions, and introduce the (nondeterministic) class NLD of decision problems for which there exists a certificate that can be verified in constant number of communication rounds. We prove that there exists an NLD-complete problem. We also show that there exist problems not in NLD. On the other hand, we prove that the class NLD#n, which is NLD assuming that each processor can access an oracle that provides the number of nodes in the network, contains all (decidable) languages. For this class we provide a natural complete problem as well.

5.3.3. Locality and checkability in wait-free computing

Participants: Pierre Fraigniaud, Sergio Rajsbaum, Travers Corentin.

The paper [9],studies notions of locality that are inherent to the specification of distributed tasks, and independent of the computing model, by identifying fundamental relationships between the various scales of computation, from the individual process to the whole system. A locality property called *projection-closed* is identified. This property completely characterizes tasks that are wait-free *checkable*, where a task $T = (\mathcal{I}, \mathcal{O}, \Delta)$ is said to be checkable if there exists a distributed algorithm that, given $s \in \mathcal{I}$ and $t \in \mathcal{O}$, determines whether $t \in \Delta(s)$, i.e., whether t is a valid output for s according to the specification of T. Projection-closed tasks are proved to form a rich class of tasks. In particular, determining whether a projection-closed task is wait-free solvable is shown to be undecidable. A stronger notion of locality is identified by considering tasks whose outputs "look identical" to the inputs at every process: a task $T = (\mathcal{I}, \mathcal{O}, \Delta)$ is said to be *locality-preserving* if $\mathcal{O}$ is a covering complex of $\mathcal{I}$. We show that this topological property yields obstacles for wait-free solvability different in nature from the classical impossibility results. On the other hand, locality-preserving tasks are projection-closed, and thus they are wait-free checkable. A classification of locality-preserving tasks in term of their relative computational power is provided. This is achieved by defining a correspondence between subgroups of the *edgepath* group of an input complex and locality-preserving tasks. This correspondence enables to demonstrate the existence of hierarchies of locality-preserving tasks, each one containing, at the top, the universal task (induced by the universal covering complex), and, at the bottom, the trivial identity task.

5.3.4. Delays Induce an Exponential Memory Gap for Rendezvous in Trees

Participants: Pierre Fraigniaud, Pelc Andrzej.

The aim of rendezvous in a graph is meeting of two mobile agents at some node of an unknown anonymous connected graph. In this paper [8], we focus on rendezvous in trees, and, analogously to the efforts that have been made for solving the exploration problem with compact automata, we study the size of memory of mobile agents that permits to solve the rendezvous problem deterministically. We assume that the agents are identical, and move in synchronous rounds. We first show that if the delay between the starting times of the agents is *arbitrary*, then the lower bound on memory required for rendezvous is $\Omega(\log n)$ bits, even for the line of length n. This lower bound meets a previously known upper bound of $O(\log n)$ bits for rendezvous in arbitrary graphs of size at most n. Our main result is a proof that the amount of memory needed for rendezvous *with simultaneous start* depends essentially on the number ℓ of leaves of the tree, and is exponentially less impacted by the number n of nodes. Indeed, we present two identical agents with $O(\log \ell + \log\log n)$ bits of memory that solve the rendezvous problem in all trees with at most n nodes and at most ℓ leaves. Hence, for the class of trees with polylogarithmically many leaves, there is an exponential gap in minimum memory size needed for rendezvous between the scenario with arbitrary delay and the scenario with delay zero. Moreover, we show that our upper bound is optimal by proving that $\Omega(\log \ell + \log\log n)$ bits of memory are required for rendezvous, even in the class of trees with degrees bounded by 3.

5.3.5. On the Manipulability of Voting Systems: Application to Multi-Operator Networks

Participants: François Durand, Fabien Mathieu, Ludovic Noirie.

Internet is a large-scale and highly competitive economic ecosystem. In order to make fair decisions, while preventing the economic actors from manipulating the natural outcome of the decision process, game theory is a natural framework, and voting systems represent an interesting alternative that, to our knowledge, has not yet being considered. They allow competing entities to decide among different options. In this paper [20], we investigate their use for end-to-end path selection in multi-operator networks, analyzing their manipulability by tactical voting and their economic efficiency.We show that Instant Runoff Voting is much more efficient and resistant to tactical voting than the natural system which tries to get the economic optimum.

5.4. Communication and Fault Tolerance in Distributed Networks

5.4.1. Linear Space Bootstrap Communication Schemes

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Sergio Rajsbaum.

We consider in [18], a system of n processes with ids not a priori known, that are drawn from a large space, potentially unbounded. How can these n processes communicate to solve a task? We show that n a priori allocated Multi-Writer Multi-Reader (MWMR) registers are both needed and sufficient to solve any read-write wait free solvable task. This contrasts with the existing possible solution borrowed from adaptive algorithms that require $\Theta(n^2)$ MWMR registers. To obtain these results, the paper shows how the processes can non blocking emulate a system of n Single-Writer Multi-Reader (SWMR) registers on top of n MWMR registers. It is impossible to do such an emulation with $n-1$ MWMR registers. Furthermore, we want to solve a sequence of tasks (potentially infinite) that are sequentially dependent (processes need the previous task's outputs in order to proceed to the next task). A non blocking emulation might starve a process forever. By doubling the space complexity, using $2n-1$ rather than just n registers, the computation is wait free rather than non blocking.

5.4.2. Black Art: Obstruction-Free k-set Agreement with |MWMR registers| < |proccesses|

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Sergio Rajsbaum.

When n processes communicate by writing to and reading from $k < n$ MWMR registers the "communication bandwidth" precludes emulation of SWMR system, even non-blocking.

Nevertheless, recently a positive result was shown that such a system either wait-free or obstruction-free can solve an interesting one-shot task. This paper demonstrates another such result. It shows that $(n-1)-$set agreement can be solved obstruction-free with merely 2 MWMR registers. Achieving $k-$set agreement with $n-k+1$ registers is a challenge. In [17], we make the first step toward it by showing $k-$set agreement with $2(n-k)$ registers.

5.4.3. Adaptive Register Allocation with a Linear Number of Registers

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Leslie Lamport.

In [16], we give an adaptive algorithm in which processes use multi-writer multi- reader registers to acquire exclusive write access to their own single-writer, multi-reader registers. It is the first such algorithm that uses a number of registers linear in the number of participating processes. Previous adaptive algorithms require at least $\Theta(n^{3/2})$ registers

5.4.4. Uniform Consensus with Homonyms and Omission Failures

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

In synchronous message passing models in which some processes may be homonyms, i.e. may share the same id, we consider the consensus problem. Many results have already been proved concerning Byzantine failures in models with homonyms, we complete in [19], the picture with crash and omission failures.

Let n be the number of processes, t the number of processes that may be faulty ($t < n$) and $l(1 \leq l \leq n)$ the number of identifiers. We prove that for crash failures and send-omission failures, uniform consensus is solvable even if $l = 1$, that is with fully anonymous processes for any number of faulty processes.

Concerning omission failures, when the processes are numerate, i.e. are able to count the number of copies of identical messages they received in each round, uniform consensus is solvable even for fully anonymous processes for $n > 2t$. If processes are not numerate, uniform consensus is solvable if and only if $l > 2t$.

All the proposed protocols are optimal both in the number of communication steps needed, and in the number of processes that can be faulty.

All these results show, (1) that identifiers are not useful for crash and send-omission failures or when processes are numerate, (2) for general omission or for Byzantine failures the number of different ids becomes significant.

5.4.5. Byzantine agreement with homonyms

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Rachid Guerraoui, Anne-Marie Kermarrec, Hung Tran-The.

So far, the distributed computing community has either as- sumed that all the processes of a distributed system have distinct identifiers or, more rarely, that the processes are anonymous and have no identifiers. These are two extremes of the same general model: namely, n processes use l dif- ferent authenticated identifiers, where $1 \leq l \leq n$. In this paper [3], we ask how many identifiers are actually needed to reach agreement in a distributed system with t Byzantine processes. We show that having $3t + 1$ identifiers is necessary and sufficient for agreement in the synchronous case but, more sur- prisingly, the number of identifiers must be greater than $(n + 3t)/2$ in the partially synchronous case. This demonstrates two differences from the classical model (which has l = n): there are situations where relaxing synchrony to partial syn- chrony renders agreement impossible; and, in the partially synchronous case, increasing the number of correct processes can actually make it harder to reach agreement. The im- possibility proofs use the fact that a Byzantine process can send multiple messages to the same recipient in a round. We show that removing this ability makes agreement easier: then, $t + 1$ identifiers are sufficient for agreement, even in the partially synchronous model.

5.4.6. Byzantine agreement with homonyms in synchronous systems

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

We consider in [4], the Byzantine agreement problem in synchronous systems with homonyms. In this model different processes may have the same authenticated identifier. In such a system of n processes sharing a set of l identifiers, we define a distribution of the identifiers as an integer partition of n into l parts $n_1 ..., n_l$ giving for each identifier i the number of processes having this identifier.

Assuming that the processes know the distribution of identifiers we give a necessary and sufficient condition on the integer partition of n to solve the Byzantine agreement with at most t Byzantine processes. Moreover we prove that there exists a distribution of l identifiers enabling to solve Byzantine agreement with at most t Byzantine processes if and only if $n > 3t$, $l > t$ and $l \frac{(n-r)t}{n-t-\min(t,r)}$ where $r = n \mod l$.

This bound is to be compared with the $l > 3t$ bound proved in Delporte-Gallet et al. (2011) when the processes do not know the distribution of identifiers.

5.4.7. *Convergence of the D-iteration algorithm: convergence rate and asynchronous distributed scheme*

Participants: Dohy Hong, Fabien Mathieu, Gérard Burnside.

In this paper [25], we define the general framework to describe the diffusion operators associated to a positive matrix. We define the equations associated to diffusion operators and present some general properties of their state vectors. We show how this can be applied to prove and improve the convergence of a fixed point problem associated to the matrix iteration scheme, including for distributed computation framework. The approach can be understood as a decomposition of the matrix-vector product operation in elementary operations at the vector entry level.

5.5. Discrete Optimization Algorithms

5.5.1. *Shrinking Maxima, Decreasing Costs: New Online Packing and Covering Problems*

Participants: Pierre Fraigniaud, Magnús M. Halldórsson, Boaz Patt-Shamir, Dror Rawitz, Adi Rosén.

We consider in [23], two new variants of online integer programs that are duals. In the packing problem we are given a set of items and a collection of knapsack constraints over these items that are revealed over time in an online fashion. Upon arrival of a constraint we may need to remove several items (irrevocably) so as to maintain feasibility of the solution. Hence, the set of packed items becomes smaller over time. The goal is to maximize the number, or value, of packed items. The problem originates from a buffer-overflow model in communication networks, where items represent information units broken into multiple packets. The other problem considered is online covering: There is a universe to be covered. Sets arrive online, and we must decide for each set whether we add it to the cover or give it up. The cost of a solution is the total cost of sets taken, plus a penalty for each uncovered element. The number of sets in the solution grows over time, but its cost goes down. This problem is motivated by team formation, where the universe consists of skills, and sets represent candidates we may hire. The packing problem was introduced for the special case where the matrix is binary; in this paper we extend the solution to general matrices with non-negative integer entries. The covering problem is introduced in this paper; we present matching upper and lower bounds on its competitive ratio.

5.5.2. *Generalized Subdifferentials of the Sign Change Counting Function*

Participants: Dominique Fortin, Ider Tseveendorj.

A natural generalization of piecewise linear approximation of non convex problems relies on piecewise convex approximation; along the way to solve the piecewise convex maximization problem [30] both effectively and efficiently, optimality conditions have to be addressed in two ways: either the violation of necessary conditions should lead to a direction of improvement from a local solution, or a sufficient condition for global optimality has to be fulfilled. The way to either goal is paved with subdifferentials and their generalizations on a per problem basis.

In the article [29], the counting function on binary values is extended to the signed case in order to count the number of transitions between contiguous locations. A generalized subdifferential for the sign change counting function is given where classical subdifferentials remain intractable. An attempt to prove global optimality at some point, for the 4-dimensional first non trivial example, is made by using a sufficient condition specially tailored among all the cases for this subdifferential.

HIPERCOM2 Team

6. New Results

6.1. Wireless Sensor Networks

6.1.1. Node activity scheduling and routing in Wireless Sensor Networks

Participants: Cédric Adjih, Ichrak Amdouni, Pascale Minet.

The need to maximize network lifetime in wireless ad hoc networks and especially in wireless sensor networks requires the use of energy efficient algorithms and protocols. Motivated by the fact that a node consumes the least energy when its radio is in sleep state, we achieve energy efficiency by scheduling nodes activity. Nodes are assigned time slots during which they can transmit and they can turn off their radio when they are neither transmitting nor receiving. Compared to classical TDMA-based medium access scheme, spatial bandwidth use is optimized: non interfering nodes are able to share the same time slots, collisions are avoided and overhearing and interferences are reduced. In our work about time slots assignment, two cases are studied. First, when nodes require equal channel access, we use node coloring. Second, when nodes have heterogeneous traffic demands, we designed the traffic aware time slot assignment algorithm TRASA. Unlike the majority of previous works, we generalize the definition of node coloring and slot allocation problems. Indeed, we set the maximum distance between two interfering nodes as a parameter of these problems. We prove that they are NP-complete, making heuristic approaches inevitable in practice. A central directive of this thesis is to design self-adaptive solutions. This adaptivity concerns many aspects such as the mission given by the application, the heterogeneity of node traffic demands, the network density, the regularity of network topology, and the failure of wireless links.

In the GETRF project, we target the energy efficiency in wireless sensor networks. We proposed node activity scheduling approaches that determine active and inactive slots for sensor nodes as to enable them to turn off their radio and save energy in the inactive slots.

1. First, we proposed a scheduling algorithm based on node coloring of grid sensor networks called VCM. This proposal was strengthen with mathematical analysis of the optimal number of colors needed to color an infinite grid. VCM produced an optimal number of colors when the transmission range tends to infinity. Also, this algorithm does not require message exchange between sensors to determine colors.

2. Second, this work was extended to adapt it to general graphs: the graph is divided into cells and the color of the cell is the color of the node on the left bottom of the cell. Nodes inside the cell are scheduled successively.

In addition to the energy efficiency, we targeted the delay optimization for data collection applications in grid wireless sensor networks. We profit from the previous work VCM and integrate it with a new hierarchical routing method to minimize data collection delays.

6.1.2. Time slot and channel assignment in multichannel Wireless Sensor Networks

Participants: Pascale Minet, Ridha Soua, Erwan Livolant.

Applying WSNs in industrial environment requires fast and reliable data gathering (or data convergecast). If packets are forwarded individually to the sink, it is called raw data convergecast. We resort to the multichannel paradigm to enhance the data gathering delay, the robustness against interferences and the throughput. Since some appli- cations require deterministic and bounded convergecast delays, we target conflict free joint time slot and channel assignment solutions that minimize the schedule length. Such solutions allow nodes to save energy by sleeping in any slot where they are not involved in transmissions. We extend existing multichannel results to take into account a sink equipped with multiple radio interfaces and heterogeneous traffic demands. Indeed, we compute the theoretical bounds, that is the minimum number of time slots needed to complete convergecast, in various topologies with different traffic demands. These bounds are provided for different acknowledgment policies. For each of them, we provide a graph-based interference model. We also give optimal schedules that achieve these optimal bounds. We formalize the problem of multichannel slot assignment using integer linear programming and solve with GLPK tool for small configurations.

We propose MODESA, a centralized joint time slot and channel assignment algorithm. We prove the optimality of MODESA in lines, multilines and balanced trees topologies. By simulations, we show that MODESA outperforms TMCP, a well known subtree-based scheduling. We improve MODESA with different channel allocation strategies depending on the channel selection criteria (channels load balancing or preference of channels with the best qualities). Moreover, we show that resorting to multipath routing minimizes the convergecast delay. This work is extended in MUSIKA to take into account multi-sinks WSNs and traffic differentiation: the problem is formalized using integer linear programming and solved with GLPK. Simulations results show that the schedule length is minimized and the buffer size is reduced. We then address the adaptivity challenge. The slot assignment should be more flexible and able to adapt to application and environment variability (e.g., alarms, temporary additional demands). Theoretical bounds on the number of additional slots intro- duced to cope with traffic changes, are given. AMSA, an incremental solution, is proposed. Its performances are evaluated in two cases: retransmissions or temporary changes in appli- cation needs.

6.1.3. WSN Redeployment

Participants: Pascale Minet, Saoucene Ridene, Ines Khoufi.

This is a joint work with Telecom SudParis: Anis Laouiti.
In many applications (e.g military, environment monitoring), wireless sensors are randomly deployed in a given area. Unfortunately, this deployment is not efficient enough to ensure full area coverage and total network connectivity. Hence, all the considered area must be covered by sensors ensuring that any event is detected in the sensing range of at least one sensor. In addition, the sensor network must be connected in terms of radio communication in order to forward the detected event to the sink(s). Thus, a redeployment algorithm has to be applied in order to achieve these two goals.

In this context, we have proposed redeployment algorithms based on virtual forces. DVFA, is our Distributed Virtual Forces Algorithm. Each node in the network executes DVFA and computes its new position based on information collected from its neighbors. Performance evaluation shows that DVFA gives very good coverage rate (between 98% and 100%) and ensures the connectivity between sensors.

Moreover, in a real environment, obstacles such as trees, walls and buildings may exist and they may impact the deployment of wireless sensors. Obstacles can prohibit the network connectivity between nodes and create some uncovered holes or some accumulation of sensors in the same region. Consequently, an efficient wireless sensors deployment algorithm is required to ensure both coverage and network connectivity in the presence of obstacles. We have focused on this problem and enhanced our Distributed Virtual Force Algorithm (DVFA) to cope with obstacles. Simulation results show that DVFA gives very good performances even in the presence of obstacles.

6.1.4. Opportunistic routing cross-layer schemes for low duty-cycle wireless sensor networks

Participants: Mohamed Zayani, Paul Muhlethaler.

This is a joint work with Nadjib Aitsaadi from University of Paris 12.
The opportunistic aspect of routing is suitable with such networks where the topology is dynamic and protocols based on topological information become inefficient. Previous work initiated by Paul Muhlethaler and Nadjib Aitsaadi consisted in a geographical receiver-oriented scheme based on RI-MAC protocol (Receiver-Initiated MAC). This scheme is revised and a new contribution proposes to address the same problem with a sender-oriented approach. After scrutinising different protocols belonging to this classification, the B-MAC protocol is chosen to build a new opportunistic cross-layer scheme. Our choice is motivated by the ability of this protocol to provide to a sender the closest neighbor to the destination (typically a sink). In other words, such a scheme enables us to obtain shorter paths in terms of hops which would increase the efficiency of information delivery. In counterparts, as it relies on long preambles (property of B-MAC) to solicit all the neighborhood, it needs larger delays and energy consumption (1% of active time). Nevertheless, this proposal remains interesting as the studied networks are dedicated to infrequent event detection and are not real time-oriented.

Starting from a simulator coded by Nadjib Aitsaadi for the receiver-oriented scheme, the new scheme has been coded under many variants. On top of ideal techniques, a realistic variant has been considered and modelled. Its particularity can be summarized in the election process of the next hop. Indeed, it is based on sending bursts by the potential candidates to receive a packet from a sender. These bursts express the closeness of each candidate to the destination and correspond to the binary complement of the distance to this destination.

The opportunistic cross-layer scheme, when designed with RI-MAC, has shown solid performances in carrying the information about a rare event detection to a sink. This is verified for an event detected by several nodes. Nevertheless, the efficiency of such a design becomes less obvious when the detection is performed by a very small number of nodes. the opportunistic routing using RI-MAC relies on a minor set of potential candidates to forward a packet. In other words, a sender can only select an awake neighbor (typically closer to the sink) as the next hop. To overcome this limitation, we initially proposed to limit the number of hops to reach the sink.The principle of B-MAC perfectly matches with this idea. It is also important to highlight the ability of an opportunistic cross-layer built over B-MAC to avoid collisions. B-MAC- and RI-MAC-based proposals are suitable to convey emergency packets in dense and large WSNs when the event is reported by a significant set of nodes. When this set is limited, the sake of efficiency rather suggests a scheme based on B-MAC. It should be remembered that the proposed schemes extremely limit the energy consumption compared to classical networks.

6.1.5. Data dissemination in Urban Environment

Participants: Belhaoua Asma, Nadjib Achir, Paul Muhlethaler.

Over the last decade, wireless sensor networks have brought valid solutions to real-world monitoring problems. Sensors are now incorporated in all our modern life facilities, such as mobile phones, vehicles, buses, bus stations, bikes, etc. For example, mobile phones, with their increasing capabilities are used as voice communication device but also as a sensing device able to collect data such as image, audio, GPS position, speed, etc. All these sensors could play an important role in the provisioning of a multitude of dynamic information about their environmental trends. Considering that, WSN could be considered as a valid solution to urban monitoring problems by bringing new services for the city or for the citizens. According to the last requirement, the main question that we need to answer is how the data could be collected and/or transmitted? Several algorithms were developed recently for sensor data gathering in WSN. However, the majority of existing works on WSN has focused only on specific areas applications, such as environmental monitoring, military target tracking, weather forecast, home automation, intrusion detection, etc. In this training we studied the existing strategies of dissemination in Delay/ Disruption Tolerant Networks (DTN). The main objective is to identify those that can be applied to urban environments. We implemented and tested several strategies in the WSNet network simulator on a dense network.

6.2. Cognitive Radio Networks

6.2.1. Multichannel time slot assignment in Cognitive Radio Sensor Networks

Participants: Ons Mabrouk, Pascale Minet, Ridha Soua, Ichrak Amdouni.

This is a joint work with Hanen Idoudi and Leila Saidane from ENSI, Tunisia.
Current Wireless Sensor Networks (WSNs) are deployed over unlicensed frequency bands that face an increased level of interference from various wireless systems. Cognitive Radio Sensor Networks (CRSNs) overcome this problem by allowing sensor nodes to access new spectrum bands to minimize interferences. In this paper, we focus on the MultiChannel Time Slot Assignment problem (MC-TSA) in CRSN. Each sensor node is assigned the number of time slots it needs to transfer its own data as well as the data received from its children in the rooting tree rooted at the sink without interfering with other secondary users. Besides, sensor nodes cannot transmit on a channel occupied by a primary user. Our objective is to increase the network throughput offered to sensor nodes. We start by formulating the MC-TSA problem as an Integer Linear Program where the goal is to minimize the number of slots in the schedule.We then propose an Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). We evaluate its performance in terms of number of slots and throughput.

6.2.2. Leader election in Cognitive Radio Networks

Participants: Paul Muhlethaler, Dimitrios Milioris.

This is a joint work with Philippe Jacquet from Alcatel-Lucent Bell Labs.
In this sudy we have introduced a new algorithm (green election) to achieve a dis- tributed leader election in a broadcast channel that is more efficient than the classic Part-and-Try algorithm. The algorithm has the adavantage of having a reduced overhead $log(log(N))$ rather than $log(N)$. More importantly the algorithm has the a greatly reduced energy consumption since it requires $O(N^{1/k})$ burst transmissions instead of $O(N/k)$, per election, k being a parameter depending on the physical properties of the medium of communication.

One of the applications of green election is for wireless col- lision algorithms in particular in cognitive wireless networks where the secondary network is WiFi IEEE 802.11. Since the green election is low energy consuming, it can be used as a systematic and repetitive medium access control that will naturally prevail over the WiFi CSMA scheme.

6.3. Development, implementation and distribution of the Ey-Wifi module for the NS3 simulation tool

Participants: Hana Baccouch, Cédric Adjih, Paul Muhlethaler.

Ey-Wifi module is an ns-3 module developed within the Mobsim project. Ey-Wifi stands for Elimination-Yield for WiFi networks. The main goal of Ey-Wifi is to integrate the features of the EY-NPMA channel access scheme in the ns-3 Wifi module. EY-NPMA (Elimination-Yield Non-Pre-emptive Priority Multiple Access) is a contention based protocol that has been used as the medium access scheme in HIPERLAN type 1. The main advantages of EY-NPMA are: low collision rate, more determinism and priority support. EY-NPMA is based on active signaling (black burst): a node requests access to the medium by transmitting a burst signal. More precisely, the channel access cycle comprises three phases : priority phase, elimination phase and yield phase. Compared to Wifi, EY-NPMA adds the transmission of a burst in the elimination phase: it reduces the number of nodes, that will compete in next "yield" phase (equivalent to the contention window based access of WiFi).

Furthermore, the performances of Ey-Wifi have been evaluated and compared with those of Wifi with ns-3. Distribution of Ey-Wifi module: The module and a tutorial explaining how to use it, are available at: http://hipercom.inria.fr/Ey-Wifi

6.4. Mobile ad hoc and mesh networks

6.4.1. Geographic routing and location services

Participants: Selma Boumerdassi, Pascale Minet, Paul Muhlethaler.

Thanks to its scalable nature, geographic routing is an interesting alternative to topological routing for ad-hoc networks. In fact, in order to set up such a network, each node needs to know the location of the others and location services are in charge to provide such an information.

Two kinds of location services have been provided using either a flooding or a rendez-vous, a node in the network being chosen as a server for the rendez-vous. In the scope of our research, we have proposed different mechanisms based on social groups and/or communities and studied their impact on the control traffic of various protocols. For example, based on the simulations of SLS and SFLS using NS-2, we have demonstrated that the social behaviour of nodes has a strong impact on location services and therefore that next-generation location services should take the relationships between the network users into account.

6.4.2. Optimized Broadcast Scheme for Mobile Ad hoc Networks

Participants: Ahmed Amari, Nadjib Achir, Paul Muhlethaler.

In this training we propose an optimized broadcasting mechanism, which uses very limited signaling overhead. The main objective is to select the most appropriate relay nodes according to a given cost function. Basically, after receiving a broadcast packet each potential relay node computes a binary code according to a given cost function. Then, each node starts a sequence of transmit/listen intervals following this code. In other words, each 0 corresponds to a listening interval and each 1 to a transmit interval. During this active acknowledgment signaling period, each receiver applies the following rule: if it detects a signal during any of its listening intervals, it quits the selection process, since a better relay has also captured the packet. Finally, we split the transmission range into several sectors and we propose that all the nodes within the same sector use the same CDMA orthogonal spreading codes to transmit their signals. The CDMA codes used in two different sectors are orthogonal, which guarantees that the packet is broadcast in all possible directions.

6.5. Learning for an efficient and dynamic management of network resources and services

Participants: Dana Marinca, Pascale Minet.

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 short-term, average-term and long-term 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. We intend to use and improve this prediction technique to design dynamically adaptive regret matching algorithms that will be applied to dynamically manage the resources in wireless networks, especially in sensor networks. These algorithms will allow the network to choose an optimal behavior, otherwise called a correlated equilibrium, from a defined behaviors' set. This behavior will be able to evolve in time to adapt to the network context evolution. We will focus mainly but not exclusively on applications like: the choice of communication channels depending on the predicted quality of transmission, energy efficiency, network nodes deployment, efficient routing, and intelligent switching between available technologies in a multi-technology context.

6.6. Vehicular Ad hoc NETworks (VANETs) for car merging

Participant: Paul Muhlethaler.

This is a joint work with Oyunchimeg Shagdar from the IMARA team.
Cooperative Adaptive Cruise Control (CACC) systems are intended to make driving safer and more efficient by utilizing information exchange between vehicles (V2V) and/or between vehicles and infrastructures (V2I). An important application of CACC is safe vehicle merging when vehicles join a main road, achieved by compiling information on the movement of individual main road vehicles. To support such road safety applications, the IEEE standardized the 802.11p amendment dedicated to V2V and V2I communications.

In this study, we have seek answers to the questions as to whether the IEEE 802.11p can support merging control and how the communications performance is translated into the CACC performance. We have built an analytical model of the IEEE 802.11p medium access control (MAC) for transmissions of the ETSI-standardized Cooperative Awareness Messages (CAM) and Decentralized Environmental Notification Messages (DENM) to support merging control. We have also developed a highway merging decision algorithm. Using computer simulations, packet delivery ratio (PDR), and packet inter-reception (PIR) time of IEEE 802.11p-based V2V and V2I communications and their impact on the CACC performance have been investigated. Our study has disclosed several useful insights including that PIR and throughput provide a good indication of the CACC performance, while improving PDR does not necessarily enhance the CACC performance. Moreover, thanks to its ability to reliably provide information at constant time intervals, the V2I structure offers a better support for CACC than V2V.

IMARA Project-Team

6. New Results

6.1. ABV

Participants: Hao Li, Paulo Lopes Resende, Evangeline Pollard, Joshué Pérez Rastelli, Fawzi Nashashibi.

The ABV project builds on the HAVEit philosophy (a previous IMARA project for high speed automation) by offering higher levels of automation on highways and organizing the cooperation between human and system along novel automation levels. It differs from HAVEit by focusing on congested traffic at speeds below 50 km/h and adding fully automated driving to the automation spectrum. By automatically following congested traffic, the ABV system relieves the human driver from monotonous tasks. During fully automated driving, the human driver is not required to monitor the system, but has to take over control at the end of the application zone. Real experiments on a prototype vehicle have been conducted. The experiment objective was to realize several use-cases: lane following, changing of lane, overtaking, ACC and emergency braking. All these maneuvers have been successfully conducted several times on the Satory tracks (cf. [46], [41] for more details) during the final event of the project which took place late March 2013.

6.2. Urban Autonomous Driving

Participants: Evangeline Pollard, Guillaume Tréhard, Fawzi Nashashibi.

Beyond low speed automation, IMARA is tackling a very important issue for autonomous driving on open roads, which is: dealing with intersections. In collaboration with Valeo, Imara wants to provide innovative way to safely cross any kind of intersections for an autonomous vehicle in a urban context and without communication. The goal is to deal with intersection with different shapes, (roundabout, T junctions , X junctions, *etc*.), with different rules, specific (traffic lights, main road...) or not ("priority to the right" in France), with different traffic (busy or empty).

6.3. Vehicle to pedestrian communications

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Evangeline Pollard, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality and serious injury to the vulnerable road users. While vehicle to vehicle (V2V) communications have taken much attention in the academic and industrial sectors, very limited effort has been made for vehicle to pedestrian communications. Unlike the V2V cases, where antennas are often installed on the vehicle rooftop, pedestrian's handheld device can be carried in such a way e.g, in a bag or in a pocket, which results in poor and unpredictable communications quality. In this work, we seek to an answer to the questions of whether the Wi-Fi-based V2P communications meet the requirements of the pedestrian safety application. This year, we studied the performances of the V2P communications especially receive signal strength, packet inter-arrival time, and message delivery ratio. Moreover, in order to demonstrate the feasibility of pedestrian safety supported by the V2P communications, we developed a software tool, V2ProVu, which has the functionalities of Wi-Fi based V2P communications, collision risk calculations, and hazard alarming.

6.4. Visible light communications for platooning control

Participants: Mohammad Abualhoul, Oyunchimeg Shagdar, Mohamed Marouf, Fawzi Nashashibi.

While V2V communications is requisite for platooning stability, the existing radio communications technologies suffer from poor performance in highly dense road scenarios, which are exactly to be created for platooning. Targeting this issue, we study the applicability of visible light communications (VLC) for information exchange between the platoon members [20], [35]. Because the existing studies on VLC mainly focus on indoor applications or for communications from traffic light to vehicle, the performances of VLC for V2V is not clear. In this work, we develop a complete VLC channel and noise model by taking account of the key parameters including background noise and incidence angle. Our studies show that it is feasible to achieve up to 7 meters line of sight communication range even in the presence of optical noise at significant levels and with up to 60 degree of road curvature.

6.5. ITS-G5 for road safety and efficiency applications

Participants: Oyunchimeg Shagdar, Younes Bouchaala, Mohammad Abualhoul, Manabu Tsukada, Thierry Ernst.

To support V2V and V2I communications for road safety and efficiency applications, ETSI standardized ITS-G5 technology. One of key objectives of the SCORE@F project is to study the performance of ITS-G5 in real-world scenarios and demonstrate its applicability to road safety and efficiency applications. Under the scope of the SCORE@F project, we studied the performances of ITS-G5 for both the V2V and V2I communications based on field tests and theoretical studies with emphasis on the effects of channel in combination with MAC and some parameters of car traffic [32]. An important insight achieved from the study is that in addition to the distance dependent pathloss, the signal fading and road traffic characteristics provide significant impacts on the reliability of ITS-G5.

We also study the performance of the ITS-G5 medium access control protocol for realistic autonomous driving applications especially to seek answers to the questions of whether the IEEE 802.11p can support merging control and how the communications performance is translated into that of CACC (Cooperative Adaptive Cruise Control) [33]. The study discloses several useful insights including packet inter-arrival time and throughput but not packet delivery ratio, gives good indications of the CACC performance; the V2I communications structure is preferred over the V2V structure for CACC.

Finally, we demonstrate the low latency video streaming over ITS-G5 to support platoon and reverse parking maneuvers [21].

6.6. Cooperative driving

Participants: Joshué Pérez Rastelli, Fawzi Nashashibi.

In the scope of the French project *"Co-Drive"* one task assigned to Inria was the development of a smart controller capable of driving the vehicle, allowing it to perform optimal traversal of traffic lights in order to reduce vehicle accelerations and thus the gas emissions. This controller needs remote information regarding the traffic lights' status, the distance to it and the time needed to reach it.

Three input variables, which are the traffic light times, red light, green light and the distance to interception (DTI), were defined in fuzzy logic tool [37].

Two variables are used for the traffic light (Red and Green), where each of them has defined two completely symmetrical membership functions covering all the possible inputs. In this application the time cycle of the lights are 30 seconds for green and 20 seconds for red. The values of input membership functions were defined considering these times.

The DTI membership function (see Figure 1) gives more weight to the distance when the vehicle is closer at the intersection. In this situation, the vehicle can be inside the *short* or the *middle* label, because in these cases the response has to be faster than in the case where the vehicle is in the *long* label. The cross rule base, based on driver knowledge when the vehicle is arriving to an intersection, are defined using natural language.

Some Simulations were performed to validate the controller. However, the final implementation will be presented in 2014 during the final event of Co-Drive Project.

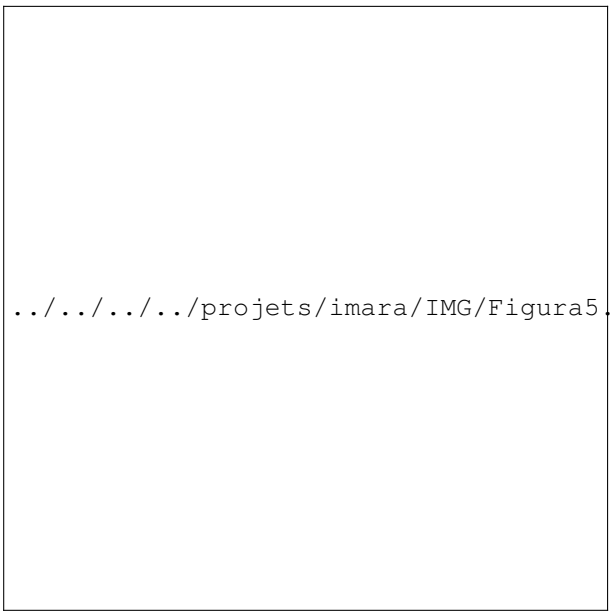


Figure 1. Codrive: input variables for the speed reference fuzzy controller

6.7. Intelligent Planning algorithm using Bezier curves

Participants: Joshué Pérez Rastelli, Fawzi Nashashibi.

The Bezier curve is the heart of the Local Planning, which allows a fast trajectory computation in order to send the trajectory in real-time to the controller stage. This method has been recently used in robot mobile solutions due to its versatility and simplicity for intersections.

We have proposed a novel method for the generation of control points for two distinct road configurations: roundabouts and a standard intersections. If an intersection is being dealt with, the control points will be generated based on the reference path given by the Global Planner.

The experiments we made presented several urban intersections. Figure 2 shows the whole generated path with four intersections and a roundabout, using the global map. A comparison with different methods is drawn. The first one (thin line) is based on the static method used in [22], which sets the control points by hand. In this case we can see how sometimes the path passes over the sidewalk. The second experiment (dotted line) is using the same previous method, but modifying the distance used to position the control points, in order to obtain a path into the road. The third method (thick line) is the Intelligent Planning algorithm. As we can see in the figure, the automatic algorithm sets the control points of Bezier (based on the convex hull property) achieving a smooth path, without going over sidewalks or obstacles.

6.8. Ontologies

Participants: Evangeline Pollard, Philippe Morignot, Fawzi Nashashibi.

Full autonomy of ground vehicles is a major goal of the ITS (Intelligent Transportation Systems) community. However, reaching such highest autonomy level in all situations (weather, traffic, . . .) is seen as impossible in practice, despite recent results regarding driverless cars (e.g., Google Cars). In addition, an automated vehicle should also self-assess its own perception abilities, and not only perceive its environment. In this new research axis, we propose an intermediate approach towards full automation, by defining a spectrum of automation

Figure 2. Generated path using different methods

levels, from fully manual (the car is driven by a driver) to fully automated (the car is driven by a computer), based on an ontological model for representing knowledge. We also propose a second ontology for situation assessment (what does the automated car perceive?), including the sensors/actuators state, environmental conditions and driver's state. Finally, we also define inference rules to link the situation assessment ontology to the automation level one [24].

6.9. Communications and Management Control for Cooperative Vehicular Systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Arnaud de La Fortelle.

One of the attractive applications of electric autonomous vehicles is electric automated Car Sharing service, where on-demand passenger transportation is provided by a set of automated vehicles and a control center, which is installed in the Internet. Data transmission from the control center to the set of vehicles requires an efficient multicast data delivery, i.e. multi-cast routing. The conventional multicast routing in the Internet is based on protocols such as Protocol Independent Multicast (PIM), which relies on a tree structure to deliver packets from the source to the destinations. Thanks to the fixed topology of the Internet, it is possible to build a large and stable multicast trees. However, due to the highly mobile nature of vehicular networks, it is not clear how stable and large can be such trees in vehicular environments. This year, we studied the stability of multicast trees for data flows from the Internet to a set of vehicles [38], [36]. Our study shows that the stability of multicast tree largely depends on the relative velocity (inter-vehicle) and the road density but not directly on the road shape or moving direction. Based on our study we are developing a mobility aware multicast routing protocol, which constructs its tree based on the vehicles' mobility dynamics and the road condition.

6.10. New urban transportation platforms: Inria's Cybus

Participants: François Charlot, Joshué Pérez Rastelli, Fawzi Nashashibi, Paulo Lopes Resende, Michel Parent, Armand Yvet.

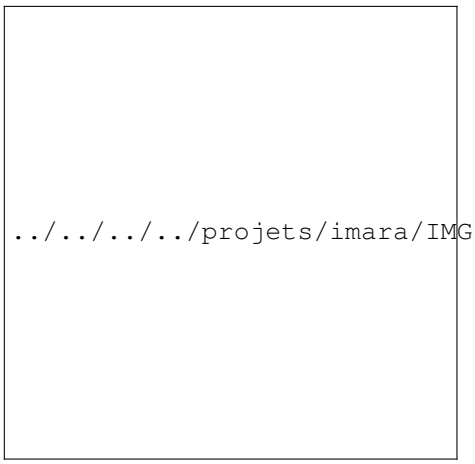


Figure 3. The Cybus operated at La Rochelle City during 3 months as a free transport service.

Cybus is the best known prototyping and demonstration platform designed at Inria. Apart from the chassis and engines, the whole hardware and software systems were developed thanks to IMARA's researchers and engineers talents. These electric vehicles are based on a Yamaha chassis but the embedded intelligence is the result of two years of development.

Much of the perception and control software has been improved. New guidance functionalities were developed this year, mainly with the introduction of stereovision-based SLAM, and Bezier curve in path planning generation. The platforms developed here (Cybus) will be demonstrated in the context of the EU CityMobil-2 project. This time real operational mobility services demonstrations will be extended to 6-12 months in selected European cities! Other showcases are expected to take place in Asian cities in 2014.

6.11. Real-time visual perception: detection and localization of static and moving objects from a moving stereo rig

Participants: Benjamin Lefaudeux, Fawzi Nashashibi.

Perception of the surrounding environment is one of the many tasks an automated vehicle has to achieve in complex and ever-changing surroundings. This, typically includes several distinct sub-tasks, such as map-building, localization, static obstacles and moving objects detection and identification. Some of these tasks are nowadays very well known, such as the map-building process which has been extensively investigated in the last decade ; whereas the perception, localization and classification of moving objects from an equally moving vehicle are in many aspects a work in progress. The objective of the PhD thesis of Benjamin Lefaudeux was to propose a vision-based approach built on the extensive tracking of numerous visual features over time, from a stereo-vision pair.

Through on-the-fly environment 3D reconstruction, based on visual clues, we proposed an integrated method to detect and localize static and moving obstacles, whose position, orientation and speed vector is estimated. Our implementation runs in real-time depending on the number of processed points, and should in the future be enclosed in a more complete, probabilistic pipeline. The complete achievements are described in the thesis of Benjamin Lefaudeux ([8] defended on September 30th) with very interesting and competitive results obtained with international benchmarks (cf. Figure 4) and on the real vehicles of IMARA.

6.12. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Victorin Martin.

Figure 4. Left: A single camera view from the KITTI sequence. Right: A bird view of the scene as modeled by the system: point cloud and estimated trajectory.

This work [55] 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's highlights are

- Victorin Martin has defended his PhD thesis entitled "Modélisation probabiliste et inférence par l'algorithme Belief Propagation" [9] at Mines-ParisTech on May 23.
- The work about the theoretical aspects of encoding real valued variables into a binary Ising model has been published as a research report [44] and submitted for publication.

6.13. Sparse covariance inverse estimate for Gaussian Markov Random Field

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Victorin Martin.

We investigate the problem of Gaussian Markov random field (GMRF) selection under the constraint that the model is suitable for Gaussian belief propagation (GaBP) inference. We develop a method based on iterative proportional scaling (IPS) to incrementally select optimal GMRF factors, while maintaining GaBP compatibility. Besides the intrinsic sparsity-inducing capability, the proposed method is indeed sufficiently flexible to incorporate various spectral constraints like e.g. walk summability (WS) to insure the compatibility of the solutions with Gaussian Belief Propagation inference. Experimental tests on various datasets with refined L_0 or L_1 regularized sparse inverse estimate indicate that this approach is competitive and provides us with useful alternatives to traditional sparsity-inducing penalizations norms, giving more freedom in the graph structure selection process with no additional computational cost.

6.14. Evaluation of dual mode transport system by event-driven simulation

Participants: Arnaud de La Fortelle, Jean-Marc Lasgouttes, Thomas Liennard.

The European project CATS — City Alternative Transport System — is developing and evaluating a new vehicle system using a single type of vehicle for two different usages: individual use or collective transport. Real experiments will necessarily take place with a limited number of vehicles and stations. Hence, there is a need for evaluation using simulations.

We are developing a discrete events simulator for that purpose, which model relies on an adapted events/decision graph. The new feature of this model is the way we deal with two modes that can be extended to many other modes. This work therefore shows on a concrete example a method to efficiently merge multiple modes into one model.

This year has seen a partial rewrite of the simulator in order to make it more generic and handle the new setting of the CATS project with automated vehicles.

6.15. Herding behavior in a social game

Participants: Guy Fayolle, Jean-Marc Lasgouttes.

The system *Ma Micro Planète* belongs to the so-called *Massively Multi-Player online Role Playing game* (MMORPG), its main goal being to incite users to have a sustainable mobility. Two objectives have been pursued.

- Construct an experimental platform to collect data in order to prompt actors of the mobility to share information (open data system).
- See how various mechanisms of a game having an additive effect could modify the transportation requests.

At the heart of the game are community-driven *points of interest* (POIs), or *sites*, which have a score that depends on the players activity. The aim of this work is to understand the dynamics of the underlying stochastic process. We analyze in detail its stationary regime in the thermodynamic limit, when the number of players tends to infinity. In particular, for some classes of input sequences and selection policies, we provide necessary and sufficient conditions for the existence of a complete meanfield-like measure, showing off an interesting *condensation* phenomenon.

The work has been published this year in *Queueing Systems* [11].

6.16. Analytic properties of random walks in the quarter plane

Participant: Guy Fayolle.

In collaboration with K. Raschel (CNRS, Université F. Rabelais à Tours), we pursued the works initiated these last three years in two main directions.

6.16.1. The group and zero drift case

In several recent studies on random walks with small jumps in the quarter plane, it has been noticed that the so-called *group of the walk* governs the behavior of a number of quantities, in particular through its *order*. When the *drift* of the random walk is equal to 0, we have provided an effective criterion (see RA 2012) giving the order of this group. More generally, we showed that in all cases where the *genus* of the algebraic curve defined by the so-called *kernel* is 0, the group is infinite, except precisely for the zero drift case, where finiteness is quite possible.

This year, we investigated new proofs of this results, which could lead to an explicit tractable criterion for the finiteness of the group, which a priori, as shown in [2] involves a ratio of elliptic integrals.

6.16.2. Counting and asymptotics

The enumeration of planar lattice walks is a classical topic in combinatorics. For a given set $\mathcal{S}$ of allowed unit jumps (or steps), it is a matter of *counting the number of paths* starting from some point and ending at some arbitrary point in a given time, and possibly restricted to some regions of the plane.

Like in the probabilistic context, a common way of attacking these problems relies on the following analytic approach. Let $f(i,j,k)$ denote the number of paths in $\mathbb{Z}_+^2$ starting from $(0,0)$ and ending at (i,j) at time k. In the case of small jumps (size at most one), the corresponding CGF

$$F(x,y,z) = \sum_{i,j,k\geq 0} f(i,j,k)x^i y^j z^k$$

satisfies the functional equation

$$K(x,y,z)F(x,y,z) = c(x)F(x,0,z) + \widetilde{c}(y)F(0,y,z) + c_0(x,y),$$

where x,y,z are complex variables, $K(x,y,z)$ is a polynomial of degree 2 (both in x and y), and linear in the time variable z which plays somehow the role of a parameter. The question of the type of the associated counting generating functions, rational, algebraic, or holonomic (i.e. solution of a linear differential equation with polynomial coefficients), was solved whenever the group is *finite* (see RA 2010). When the group is infinite, the problem is still largely open.

The nature of the singularities of the function F plays a key role for this classification. Starting from our study [54], we proved in various cases that the first singularities of F(1,0,z) are either polar or correspond to a value z_g for which the genus of the algebraic curve $K(x,y,z) = 0$ passes from 1 to 0 (i.e. a torus becomes a sphere).

6.16.3. Harmonic functions and more general jumps

The determination of Martin boundaries in the case of random walks is a longstanding problem, solved only in special situations. For homogeneous random walks in the quarter plane, stopped on the boundary (the axes), with upward jumps of size 1, and arbitrary downward jumps of size d, it turns out that the computation of harmonic functions is here plainly equivalent to find a positive function H satisfying a functional equation of the form

$$L(x,y)H(x,y) = L(x,0)H(x,0) + L(0,y)H(0,y) - L(0,0)H(0,0).$$

Here the chief difficulty to make the reduction to a boundary value problem is to analyze the algebraic curve $L(x,y) = 0$, which might be of arbitrary genus. Some examples lead us to conjecture the existence of a *single real cut* inside the unit disk, which should allow to get integral form solution.

6.16.4. Correction of papers

Guy Fayolle found important errors in several articles dealing with models involving random walks in the quarter plane. This is the object of the letter to the editors [10]. The Concerned authors are currently preparing corrected versions.

MATHRISK Project-Team

6. New Results

6.1. Credit risk

Participants: Aurélien Alfonsi, Céline Labart, Jérôme Lelong.

We have ended our study on stochastic local intensity model. We have shown by the mean of a particles system that this model is well defined and have obtained an efficient way to perform Monte-Carlo algorithms for this model.

6.2. Liquidity risk

Participants: Aurélien Alfonsi, A. Schied.

A. Alfonsi and A. Schied (Mannheim University) are working on price impact models that describe how the price is modified by large trades. The paper together with J. Acevedo on a time-dependent price impact is now accepted for publication. With A. Schied and F. Klöck [45], we have studied the cross price impact between different assets and identified conditions on the resilience of this impact that avoid manipulations strategies. With P. Blanc, we are working on the optimal execution problem when there are many large traders that modify the price.

6.3. Systemic Risk

Participants: Andreea Minca, Agnès Sulem.

We are working on the theory of the stochastic control of financial networks.

In two related articles, we find the optimal strategy of a government who seeks to make equity infusions in a banking system prone to insolvency and to bank runs. The first article combines stochastic control and the random graph representation of the financial system developed in Andreea's thesis. The second article combines the network representation of a financial system and the solvency-based mechanism of contagion with another potent source of distress, which is funding illiquidity [31] and [60].

6.4. Estimation of the parameters of a Wishart process

Participants: Aurélien Alfonsi, Ahmed Kebaier, Clément Rey.

This research has started this year together with the thesis of Clément Rey. We are studying the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and the non ergodic case.

6.5. An Affine term structure model for interest rates that involve Wishart diffusions

Participants: Aurélien Alfonsi, E. Palidda.

Affine term structure models (Dai and Singleton, Duffie, ...) consider vector affine diffusions. Here, we would like to extend this model by including some Wishart dynamics, and to get a model that could better fit the market. We also develop some numerical pricing methods for this model to make its implementation possible.

6.6. Applications of optimal transport

Participants: Aurélien Alfonsi, Benjamin Jourdain, Arturo Kohatsu-Higa.

A. Alfonsi and B. Jourdain study the Wasserstein distance between two probability measures in dimension n sharing the same copula C. The image of the probability measure dC by the vectors of pseudo-inverses of marginal distributions is a natural generalization of the coupling known to be optimal in dimension $n = 1$. In dimension $n > 1$, it turns out that for cost functions equal to the p-th power of the L^q norm, this coupling is optimal only when $p = q$ i.e. when the cost function may be decomposed as the sum of coordinate-wise costs.

As another application of optimal transport, they are working with A. Kohatsu-Higa on the uniform in time estimation of the Wasserstein distance between the time-marginals of an elliptic diffusion and its Euler scheme. To generalize in higher dimension the result that they obtained previously in dimension one using the optimality of the explicit inverse transform, they compute the derivative of the Wasserstein distance with respect to the time variable thanks to the theory developped by Ambrosio Gigli and Savare. The abstract properties of the optimal coupling between the time marginals then enable them to estimate this time derivative.

6.7. Capital distribution and portfolio performance in the mean-field Atlas model

Participants: Benjamin Jourdain, J. Reygner.

B. Jourdain and J. Reygner study a mean-field version of rank-based models of equity markets, introduced by Fernholz in the framework of stochastic portfolio theory. They first obtain an asymptotic description of the market when the number of companies grows to infinity. They then discuss the long-term capital distribution in this asymptotic model, as well as the performance of simple portfolio rules. In particular, they highlight the influence of the volatility structure of the model on the growth rates of portfolios.

6.8. Public Private Partnerships

Participants: Gilles Edouard Espinosa, Caroline Hillairet, Benjamin Jourdain, Monique Pontier.

With Gilles Edouard Espinosa, Caroline Hillairet and Monique Pontier, Benjamin Jourdain is interested in the problem of outsourcing the debt for a big investment, according two situations: either the firm outsources both the investment (and the associated debt) and the exploitation to a private consortium, or the firm supports the debt and the investment but outsources the exploitation. They prove the existence of Stackelberg and Nash equilibria between the firm and the private consortium, in both situations. They compare the benefits of these contracts.They conclude with a study of what happens in case of incomplete information, in the sense that the risk aversion coefficient of each partner may be unknown by the other partner [51].

6.9. Backward Stochastic (Partial) Differential equations with jumps and stochastic control

Participants: Roxana-Larisa Dumitrescu, Marie-Claire Quenez, Agnès Sulem.

We have studied optimization problems for BSDEs with jumps, optimal stopping for dynamic risk measures induced by BSDEs with jumps and associated reflected BSDEs, and generalized Dynkin games associated to double barriers reflected BSDEs with jumps [32], [38], [42]. A. Sulem, with B. Øksendal and T. Zhang has also studied optimal stopping for Stochastic Partial Differential equations and associated reflected SPDEs [34], and optimal control of Forward-Backward SDEs [54].

6.10. Utility maximization and Arbitrage Theory

Participants: Claudio Fontana, Bernt Øksendal, Agnès Sulem.

B. Øksendal and A. Sulem have contributed to the issue of robust utility maximization in jump diffusion markets via a stochastic maximum approach and the links with robust duality [53].

In the period January - October 2013, the main subject of investigation of C. Fontana has been arbitrage theory, with a special emphasis on no-arbitrage conditions weaker than the classical notion of No Free Lunch with Vanishing Risk (NFLVR). In particular, in the context of financial market models based on diffusion processes (see [35]), we have provided a characterization of several no-arbitrage conditions as well as a generalization of the second fundamental theorem of asset pricing. In the context of jump-diffusion models under partial information (see [25]), we have studied the relation between market viability (in the sense of solvability of portfolio optimization problems) and the existence of a martingale measure given by the marginal utility of terminal wealth, without a-priori assuming no-arbitrage restrictions on the model. Finally, in the paper [41], we have provided a critical analysis of the paper Arbitrage, Approximate Arbitrage and the Fundamental Theorem of Asset Pricing (Wong & Heyde, 2010), where the authors aim at proposing an original and simple proof of the fundamental theorem of asset pricing in the context of incomplete diffusion-based models. We have shown that the method of Wong & Heyde (2010) can only work in the well-known case of complete markets, exhibiting an explicit counterexample.

6.11. Regularity of probability laws using an interpolation method

Participant: Vlad Bally.

This work was motivated by previous papers of Nicolas Fournier, J. Printemps, E. Clément, A. Debusche and of myself, concerning the regularity of the law of the solutions of some equations with coefficients with little regularity - for example diffusion processes with Hölder coefficients (but also many other examples including jump type equations, Bolzmann equation or Stochastic PDE's). Since we do not have sufficient regularity the usual approach by Malliavin calculus fails in this framework. Then one may use an alternative idea which roughly speaking is the following: We approximate the law of the random variable X (the solution of the equation at hand) by a sequence $X(n)$ of random variables which are smooth and consequently we are able to establish integration by parts formulas for $X(n)$ and we are able to obtain the absolutely continuity of the law of $X(n)$ and to establish estimates for the density of the law of $X(n)$ and for its derivatives. Notice that the derivatives of the densities of $X(n)$ generally blow up - so we can not derive directly results concerning the density of the law of X. But, if the speed of convergence of $X(n)$ to X is stronger than the blow up, then we may obtain results concerning the density of the law of X. It turns out that this approach fits in the framework of interpolation spaces and that the criterion of regularity for the law of X amounts to the characterization of an interpolation space between a space of distributions and a space of smooth functions. Although the theory of interpolation spaces is very well developed and one already know to characterize the interpolation spaces for Sobolev spaces of positive and negative indices, we have not found in the (huge) literature a result which covers the problem we are concerned with. So, although our result may be viewed as an interpolation result, it is a new one. The above work is treated in the paper [62] (in collaboration with Lucia Caramellino). As an application we discussed in [48] the regularity of the law of a Wiener functional under a Hörmander type non degeneracy condition.

6.12. A stochastic parametric representation for the density of a Markov process

Participant: Vlad Bally.

Classical results in the PDE theory (due to A. Friedmann) assert that, under uniform ellipticity conditions, the law of a diffusion process has a continuous density (the approach of A. Friedmann is analytical and concerns PDE's instead of the corresponding diffusion process). The method developed by A. Friedmann becomes well known as the "parametric method". In collaboration with A. Kohatzu Higa [49] we gave a probabilistic approach which represents the probabilistic counterpart of the parametric method. We obtained a probabilistic representation for the density of the law of the solution of a SDE and more generaly, for a class of Markov processes including solutions of jump type SDE's. This representation may be considered as a perfect simulation scheme and so represents a starting point for Monte Carlo simulation. However the random variable which appears in the stochastic representation has infinite variance, so direct simulation gives unstable results (as some preliminary tests have proved). In order to obtain an efficient simulation scheme some more work on the reduction of variance has to be done.

6.13. Regularity of probability laws using an interpolation method

Participant: Vlad Bally.

The distance between two density functions and convergence in total variation. In collaboration with Lucia Caramellino we obtained estimates of the distance between the densities of the law of two random variables using an abstract variant of Malliavin calculus. We used these estimates in order to study the convergence in total variation of a sequence of random variables. This has been done in [47]. We are now working on more specific examples concerning the Central Limit Theorem. In the last years the convergence in entropy distance and in total variation distance for several variants of the CLT has been considered in papers of S. Bobkov, F. Gotze, G. Peccati, Y. Nourdin, D. Nualart and G. Polly. So this seems to be a very active research area. Moreover, in an working paper in collaboration with my Phd student R. Clement, we use the same methods in order to study the total variation distance between two Markov semigroups and in particular for approximation schemes. A special interest is devoted to higher order schemes - as for example the Victoire Nyomia scheme.

MICMAC Project-Team

5. New Results

5.1. Electronic structure calculations

Participants: Eric Cancès, Ismaila Dabo, Virginie Ehrlacher, David Gontier, Salma Lahbabi, Claude Le Bris, Gabriel Stoltz.

In electronic structure calculation as in most of our scientific endeavours, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

E. Cancès and S. Lahbabi have addressed issues related to the modeling and simulation of defects in periodic crystals. Computing the energies of local defects in crystals is a major issue in quantum chemistry, materials science and nano-electronics. In collaboration with M. Lewin (CNRS, Cergy), E. Cancès and A. Deleurence have proposed in 2008 a new model for describing the electronic structure of a crystal in the presence of a local defect. This model is based on formal analogies between the Fermi sea of a perturbed crystal and the Dirac sea in Quantum Electrodynamics (QED) in the presence of an external electrostatic field. The justification of this model is obtained using a thermodynamic limit of Kohn-Sham type models. In collaboration with M. Lewin, E. Cancès and S. Lahbabi have introduced a functional setting for mean-field electronic structure models of Hartree-Fock or Kohn-Sham types for disordered quantum systems, and used these tools to study the reduced Hartree-Fock model for a disordered crystal where the nuclei are classical particles whose positions and charges are random.

D. Gontier has obtained a complete, explicit, characterization of the set of spin-polarized densities for finite molecular systems. This problem was left open in the pionnering work of von Barth and Hedin setting up the Kohn-Sham density functional theory for magnetic compounds.

On the numerical side, E. Cancès, L. He (ENPC), Y. Maday (University Paris 6) and R. Chakir (IFSTTAR) have designed and analyzed a two-grid methods for nonlinear elliptic eigenvalue problems, which can be applied, in particular, to the Kohn-Sham model. Some numerical tests demonstrating the interest of the approach have been performed with the Abinit software.

Implicit solvation models aims 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 media 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. A collaboration with F. Lipparini (Paris 6), B. Mennucci (Department of Chemistry, University of Pisa) and J.-P. Picquemal (Paris 6) is in progress to implement this algorithm in widely used computational softwares (Gaussian and Tinker), and to extend this method to other implicit solvation models.

Claude Le Bris, in collaboration with Pierre Rouchon (Ecole des Mines de Paris), 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 specific case under consideration is that of oscillation revivals of a set of atoms interacting resonantly with a slightly damped coherent quantized field of photons. The approach may be employed for other similar equations. Current work is directed towards other numerical challenges for this type of problems.

5.2. Computational Statistical Physics

Participants: Claude Le Bris, Frédéric Legoll, Tony Lelièvre, Francis Nier, Mathias Rousset, Gabriel Stoltz.

5.2.1. Free Energy calculations

For large molecular systems, the information of the whole configuration space may be summarized in a few coordinates of interest, called reaction coordinates. An important problem in chemistry or biology is to compute the effective energy felt by those reaction coordinates, called free energy.

In [39], T. Lelièvre and G. Stoltz, in collaboration with physicists from CEA Saclay (especially, M. Athenes) studied a new adaptive technique of ABF type to compute on-the-fly the free energy of a system, without evaluating the second derivatives of the reaction coordinate. The method uses a Bayesian reinterpretation of an extended system where the reaction coordinate is considered as an additional variable.

In [44], G. Fort (Telecom Paris), B. Jourdain (CERMICS), E. Kuhn (INRA), T. Lelièvre and G. Stoltz have studied the efficiency of the Wang-Landau algorithm, building on a previous study where they proved the convergence of this method. The aim was to obtain precise estimates of the exit times out of metastable states. This was done in two ways: a theoretical study in the simplest possible metastable situation, a system with three states; and a numerical study in a more realistic situation (a two-dimensional double well potential).

5.2.2. Sampling trajectories

There exist a lot of methods to sample efficiently Boltzmann-Gibbs distributions. The situation is much more intricated as far as the sampling of trajectories (and especially metastable trajectories) is concerned.

Recently, the quasi stationary distribution has been identified by the team as a good mathematical tool to analyze metastable trajectories, and to make a link between a continuous state space dynamics (Langevin dynamics) and a discrete state space dynamics (kinetic Monte Carlo models), see for examplelelievre-13. This perspective can also be used to analyze accelerated dynamics techniques which have been proposed by A. Voter in the late nineties, to simulate very efficiently the state-to-state dynamics associated with metastable trajectories. For example, in [33], T. Lelièvre with D. Aristoff (University of Minnesota) propose a mathematical analysis of the Temperature Accelerated Dynamics. In [49], T. Lelièvre and F. Nier have studied the quasi-stationary distribution for an overdamped Langevin process in a bounded domain. In the small temperature limit and by making the connection with boundary Witten Laplacians, they are able to accurately compute the spatial exit law along the boundary and non perturbative accurate formulas when the potential is changed inside the domain. This gives some insight into the foundations of the hyperdynamics method.

Finally, following a numerical observation in a previous work on the sampling of reactive trajectories by a multilevel splitting algorithm, F. Cérou (Inria Rennes), A. Guyader (Inria Rennes), T. Lelièvre and F. Malrieu (Université de Rennes) study theoretically in [19] the distribution of the lengths of these trajectories, using large deviation techniques.

5.2.3. Nonequilibrium systems

Let us also mention that the article [22] on a derivation of a Langevin-type dynamics for a heavy particle in a non-zero background flow, co-authored by M. Dobson, F. Legoll, T. Lelièvre, and G. Stoltz, has been published.

5.2.4. Sampling techniques

In [29], T. Lelièvre studies with F. Nier and G. Pavliotis (Imperial College, London) the interest of using non-reversible dynamics (overdamped Langevin dynamics with a non-gradient drift term) to efficiently sample a given Boltzmann-Gibbs distribution.

5.2.5. Numerical analysis of simulation methods

Together with B. Leimkuhler and Ch. Matthews (Edinburgh University), G. Stoltz studied in [48] the discretization errors in the computation of average properties with Langevin dynamics integrated with splitting strategies. The average properties are either static (average of a given observable) or dynamic (transport coefficients). The main tool used in this analysis is the expansion of the transition operator in powers of the time step, with exact integral remainders; as well as fine estimates on the resolvent of the Langevin operators,

especially in the so-called overdamped limit where the friction goes to infinity. Transport coefficients are studied either through errors in Green-Kubo formulae or errors in the linear response of nonequilibrium systems.

5.2.6. Coarse-graining of molecular systems

G. Stoltz, in collaboration with J.-B. Maillet and G. Faure, developed in [43] a potential energy function depending on the local density of the molecular fluid. The local density is evaluated with a three dimensional Voronoi tesselation, which proves more rigorous than the standard local averages typically found in the literature. The new potential allows to describe the compressibility of mesoparticles representing several molecules in a coarse-grained description of the atomic system. The quality of the potential has been assessed by reproducing equations of state and Hugoniot curves of model energetic materials.

5.2.7. Thermodynamic limit

The quasicontinuum method is an approach to couple an atomistic model with a coarse-grained approximation in order to compute the states of a crystalline lattice at a reduced computational cost compared to a full atomistic simulation. In that framework, the team has addressed questions related to the *finite temperature* modeling of atomistic systems and derivation of coarse-grained descriptions, such as canonical averages of observables depending only on a few variables. The work from F. Legoll and X. Blanc (Université Pierre et Marie Curie) is now published [12].

When the temperature is small, a perturbation approach can be used to compute the canonical averages of these observables depending only on a few variables, at first order with respect to temperature. The work from F. Legoll in collaboration with E. Tadmor, W. K. Kim, L. Dupuy and R. Miller on the analysis of such an approach is now also published [32].

5.2.8. Hamiltonian dynamics

Constant energy averages are often computed as long time limits of time averages along a typical trajectory of the Hamiltonian dynamics. One difficulty of such a computation is the presence of several time scales in the dynamics: the frequencies of some motions are very high (e.g. for the atomistic bond vibrations), while those of other motions are much smaller. This problem has been addressed in a two-fold manner.

Fast phenomena are often only relevant through their mean effect on the slow phenomena, and their precise description is not needed. The work from M. Dobson, C. Le Bris, and F. Legoll developing integrators for Hamiltonian systems with high frequencies (derived using homogenization techniques applied to the Hamilton-Jacobi PDE associated to the Hamiltonian ODE) is now published [22].

Another track to simulate the system for longer times is to resort to parallel computations. An algorithm in that vein is the parareal in time algorithm. The work from C. Le Bris and F. Legoll, in collaboration with X. Dai and Y. Maday, studying several variants of the original plain parareal in time algorithm, is now also published [21].

5.2.9. Effective dynamics

For a given molecular system, and a given reaction coordinate $\xi : \mathbb{R}^n \mapsto \mathbb{R}$, the free energy completely describes the statistics of $\xi(X)$ when $X \in \mathbb{R}^n$ is distributed according to the Gibbs measure. On the other hand, obtaining a correct description of the dynamics along ξ is complicated. In this context, S. Lahbabi and F. Legoll have studied in [8] the case when the fine-scale, reference dynamics is a kinetic Monte Carlo model with small and fast time scales, and proved a path-wise convergence to a coarse kinetic Monte Carlo model only retaining slow degrees of freedom.

Another question is how to use a coarse-grained description (involving only the slow degrees of freedom) as a predictor for the dynamics of the actual reference system, involving all degrees of freedom. Together with G. Samaey (KU Leuven), F. Legoll and T. Lelièvre have addressed this question in the parareal framework, and shown in [28] that the precise coupling between both models should be done carefully in order for the algorithm to be efficient. In that case, the algorithm converges to the reference full dynamics.

5.3. Complex fluids

Participants: David Benoit, Sébastien Boyaval, Claude Le Bris, Tony Lelièvre.

In his PhD under the supervision of Claude Le Bris and Tony Lelièvre, David Benoit studies models of aging fluids developed at the ESPCI (Ecole supérieure de physique et de chimie industrielles) and designed to take into account phenomena such as shear thinning, aging and shear banding in falling sphere experiments. The work consists in studying on the one hand the mathematical well-posedness of some macroscopic models, see [10], and, on the other hand, in trying to understand the link between such macroscopic models and microscopic models which have been proposed to describe such fluids, see [34].

Let us also mention that the paper [28] on a parareal algorithm to efficiently simulate micro-macro models which has been published this year.

Related to the mathematical modelling of free-surface complex flows under gravity, a new reduced model for thin layers of a viscoelastic upper-convected Maxwell fluid was derived by S. Boyaval in collaboration with François Bouchut. Possibly discontinuous solutions were numerically simulated with a new finite-volume scheme of relaxation type that satisfies a discrete counterpart of the natural dissipation [13].

This work has been pursued for other fluid models and other flow regimes, with a view to better understanding the reduction mechanism leading from a physically detailed model to a useful one for numerical simulations at large (geophysical) scales [35].

On the other hand, note that it is often possible to consider only models for *incompressible* fluids (at low Mach numbers). Now, it is both important and delicate to understand how to numerically discretize the incompressibility constraint, a long-standing issue in numerical fluid mechanics. In collaboration with M. Picasso (EPFL), S. Boyaval has thus investigated the possibility to numerically quantify *a posteriori* the quality of a well-known, "simple" numerical method discretizing the incompressibility constraint, in a simple case [36]. This is part of another effort toward useful numerical simulations of complex flows, inline with current questions focused on discretization methods..

5.4. Application of greedy algorithms

Participants: Sébastien Boyaval, Eric Cancès, Virginie Ehrlacher, Tony Lelièvre.

Model reduction techniques are very important tools for applications. They consist in deriving from a high-dimensional problem, a low-dimensional model, which gives very quickly reliable results. We are in particular interested in two techniques: Proper Generalized Decomposition (greedy algorithms) and Reduced Basis techniques.

Concerning the Proper Generalized Decomposition, current research concerns the approximation of high-dimensional spectral problems, see [38]. Prototypical applications include electronic structure calculations or the computation of buckling modes in mechanics. We also explored in the PhD of J. Infante Acevedo the application of these techniques to option pricing problems, see [45].

Finally, in [40], Fabien Casenave (CERMICS), Alexandre Ern (CERMICS), Guillaume Sylvand (EADS IW) and Tony Lelièvre propose a new non intrusive implementation of the reduced basis technique using the Empirical Interpolation Method. The interest if the method is illustrated on aeroacoustic problems.

5.5. Homogenization and related topics

Participants: Virginie Ehrlacher, Claude Le Bris, Frédéric Legoll, François Madiot, William Minvielle.

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 and P.-L. Lions, 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 endeavour is completely new. Promising results have been obtained on a simple case of a periodic system perturbed by a localized defect. Ongoing works consider other configurations, such as for instance an interface between two different crystalline phases.

The project-team also has pursued its efforts in the field of stochastic homogenization of elliptic equations, aiming at designing numerical approaches that both are pratically 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. For a given truncation of $\mathbb{R}^d$, the team has previously shown that the variance of this matrix can be reduced using the technique of antithetic variables. In [47], F. Legoll and W. Minvielle have extended this technique to nonlinear, convex homogenization problems.

In addition, F. Legoll and W. Minvielle have investigated the possibility to use other variance reduction approaches, such as control variate techniques. A promising idea is to use the weakly stochastic model previously introduced by A. Anantharaman and C. Le Bris (in which a periodic model is perturbed by a *rare* stochastic perturbation) to build a control variate model. The preliminary results that have already been obtained are very encouraging.

Yet another approach to reduce the variance is the so-called Multi Level Monte Carlo (MLMC) approach, which is based on using a surrogate model for the quantity of interest. The MLMC approach consists in using many realizations of the surrogate model (which is cheap to evaluate) and few realizations of the reference model (which is more expensive to evaluate). In collaboration with Y. Efendiev and C. Kronsbein, F. Legoll has explored in [41] how this approach can be used in random homogenization.

We have discussed above approaches to efficiently compute the homogenized coefficient, assuming we have a complete knowledge of the microstructure of the material. We have recently started to consider a related inverse problem, and more precisely a parameter fitting problem. Knowing the homogenized quantities, is it possible to recover some features of the microstructure properties? Obviously, since homogenization is an averaging procedure, not everything can be recovered from macroscopic quantities. A realistic situation is the case when we assume a functional form of the distribution of the microscopic properties, but with some unknown parameters that we would like to determine. In collaboration with A. Obliger and M. Simon, F. Legoll and W. Minvielle have started to address that problem, determining the unknown parameters of the microscopic distribution on the basis of macroscopic (e.g. homogenized) quantities. The preliminary results that have been obtained are very encouraging.

From a numerical perspective, 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 is actually still room for improvement in many different directions. In collaboration with A. Lozinski (University of Besançon), F. Legoll and C. Le Bris have introduced and studied a variant of MsFEM that considers Crouzeix-Raviart type elements on each mesh element. The continuity across edges (or facets) of the (multiscale) finite element basis set functions is enforced only weakly, using fluxes rather than point values. The approach has been analyzed (combining classical arguments from homogenization theory and finite element theory) and tested on simple, but highly convincing cases [27]. In particular, an elliptic problem set on a domain with a huge number of perforations has been considered in [37]. The variant developed outperforms all existing variants of MsFEM.

A follow up on this work, in collaboration with U. Hetmaniuk (University of Washington in Seattle) and A. Lozinski (University of Besançon), consists in the study of multiscale advection-diffusion problems. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interferes with the multiscale character of the equation is an unsolved mathematical question

worth considering for numerical purposes. This is the aim of the PhD thesis of François Madiot, which began in October 2013.

Still another question related to homogenization theory that is investigated in the group is the following. Consider an elliptic equation, say in divergence form, with a highly oscillatory matrix coefficient. Is it possible to approximate the boundary value problem for different right hand sides using a similar problem with a *constant* matrix coefficient? How can this "best" constant matrix approximating the oscillatory problem be constructed in an efficient manner? We have addressed some of these questions in [25], where we have in particular shown that this best constant matrix converges to the homogenized matrix, in the limit of infinitely rapidly oscillatory coefficients. Our approach can therefore be considered as an alternative way to compute the homogenized matrix. This is particularly interesting in random cases, where the standard approach is very expensive. Current work is directed towards extending the approach, in order to compute other quantities of interest than the homogenized coefficient.

To conclude this section, we mention the project undertaken by V. Ehrlacher during her six months postdoctoral position in the Cluster of Excellence Engineering of Advanced Materials (Erlangen University). This project, in collaboration with C. Le Bris, F. Legoll, G. Leugering and M. Stingl, aims at optimizing the shape of some materials (modelled as structurally graded linear elastic materials) in order to achieve the best mechanical response at the minimal cost. As often the case in shape optimization, the solution tends to be highly oscillatory, thus the need of homogenization techniques. We thus consider an initial microstructured material composed of steel and void and whose microstructure pattern is periodic (think e.g. of a periodic honeycomb structure). We next consider materials which are obtained from this initial material through a macroscopic deformation, and look for the optimal deformation achieving the best mechanical response. Encouraging first results have been obtained.

5.6. Coupling methods and variance reduction

Participant: Mathias Rousset.

Recently, M. Rousset has initiated a research topic on variance reduction techniques (called "asymptotic") for the simulation of stochastic models of particles. The point is to use a macroscopic (or model reduced) equation as a control variate; or in other words, to use the information of a macroscopic description to decrease the statistical error of the simulated microscopic evolution.

A first step in this program has been achieved for a microscopic model describing the individual motion of bacteriae with a Markovian velocity-jump process. The macroscopic equation is an advection-diffusion equation called the chemotaxis equation. In [30], the pobabilistic derivation of the chemotaxis equation from the individual motion of bacteriae have been carried out in a rigorous way. In [31], a numerical method simulating the individual evolution of bacteriae with "asymptotic" variance reduction have been proposed.

Motivated by the asymptotic variance reduction of DSMC methods (particle Monte-Carlo methods simulating low density fluids modeled by kinetic equations), the work in [50], M. Rousset considers space homogenous Boltzmann kinetic equations in dimension d with Maxwell collisions (and without Grad's cut-off). An explicit Markov coupling of the associated conservative (Nanbu) stochastic N-particle system is constructed, using plain parallel coupling of isotropic random walks on the sphere of two-body collisional directions. The resulting coupling is almost surely decreasing, and the L_2-coupling creation is computed explicitly. Some quasi-contractive and uniform in N coupling / coupling creation inequalities are then proved, relying on 2 + α-moments ($\alpha > 0$) of velocity distributions; upon N-uniform propagation of moments of the particle system, it yields a N-scalable α-power law trend to equilibrium. The latter are based on an original sharp inequality, which bounds from above the coupling distance of two centered and normalized random variables $(U, V) \in \mathbb{R}^d$, with the average square parallelogram area spanned by $(U - U_*, V - V_*)$, (U_*, V_*) denoting an independent copy. Two counter-examples proving the necessity of the dependance on > 2-moments and the impossibility of strict contractivity are provided. The paper, (mostly) self-contained, does not require any propagation of chaos property and uses only elementary tools.

MOKAPLAN Exploratory Action

6. New Results

6.1. Monge-Ampère solver for the Mass Transportation problem and extensions

- **Benamou, Froese (Univ. of Texas at Austin) -** We design a scheme for Aleksandrov solution of Optimal Mass Transportation between atomic measure and continuous densities. The idea is to couple the notion of viscosity solution with an adapted sub gradient discretization at dirac points where the notion of Aleksandrov solution is relevant. This would offer a "PDE" alternative to the classical gradient methods based on costly computational geometry tools [61].
- **Benamou, Collino, Mirebeau (Univ. Paris IX,CNRS) -** A new variational formulation of the determinant of a semi-definite positive matrix has been proposed based on the ideas developed in [60]. This leads to a monotone discretisation of the Monge-Ampère operator. A Newton method preserving convexity is currently being tested. The new scheme is more accurate than the wide stencil, currently the state of the art of monotone scheme for the Monge-Ampère equation.
- **Benamou, Froese (Univ. of Texas at Austin), Oberman (Univ. Mc Gill) -** When the Optimal Mass Transportation data is not balanced, i.e. the densities do not have equal mass. A natural extension of the optimal transport has been proposed by McCann and Caffareli [30] and revisited by Figalli [41]. It is formulated as an obstacle problem which automatically select the portion of mass corresponding to Optimal Mass Transportation. The numerical resolution of this problem is open and we believe ideas linked the state constraint reformulation contained in paper [6] may be applied to obtain a tractable reformulation.

6.2. Variational problems under divergence constraint - Alg2

- **Benamou, Bonne, Carlier -** Dynamic problems: we have extended the Augmented Lagrangian method used for the CFD formulation of the Optimal Mass Transportation to Mean Field Games that is for the optimal control of the continuity equation. A freefem Code has been implemented.
- **Benamou, Carlier -** Static problems with a divergence constraint. We have also extended the Augmented Lagrangian method to static problem where a space divergence constraint appears. This includes the delicate case of the original Monge Optimal Mass Transportation cost (cost=distance) and also Wardrop equilibria in congested transport and related degenerate elliptic equations, like the p-Laplacian operator. A freefem Code has been implemented.

6.3. Multi-marginal problems

- **[Carlier, Oberman (Univ. Mc Gill), Oudet (Univ. of Grenoble) -** New numerical methods for the Wasserstein barycenter and related multi-marginals problems were investigated [49]. A first method uses linear programming, in an implementation that was more efficient than expected. A second method takes advantage of the quadratic structure and leads to an efficient algorithm that can be used in texture synthesis problems arising in image processing.
- **Benamou, Carlier, Nenna** Extension of the CFD formulation and the ALG2 algorithm to the multi marginal problem with quadratic cost (Barycenter).

6.4. JKO gradient flow numerics

Figure 2. Monge transport flow between sinks and sources.

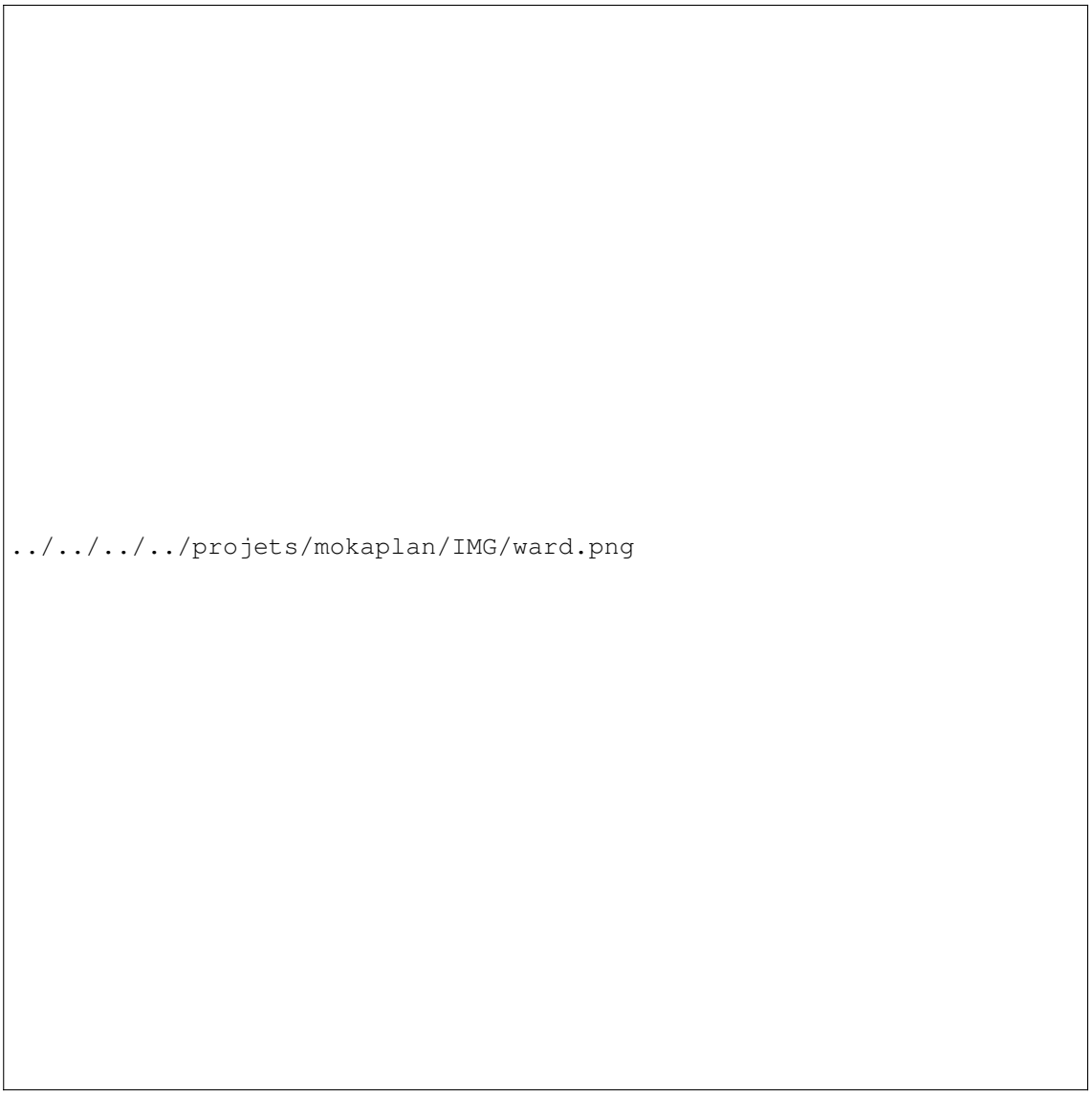


Figure 3. Congested transport flows between sink and source.

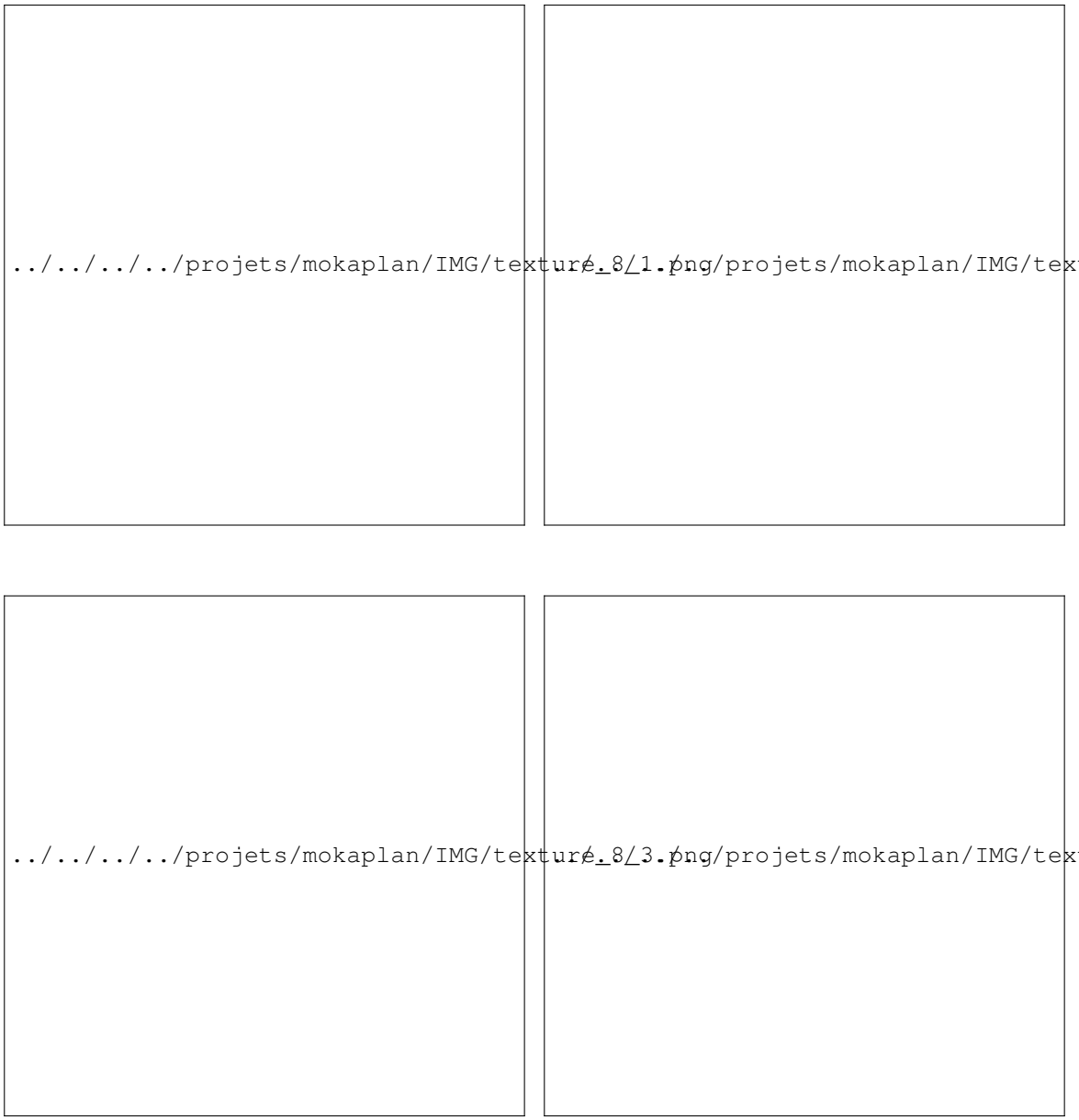


Figure 4. Texture mixing with Wasserstein barycenters, from top to bottom three densities and their barycenter.

- **Benamou, Carlier, Merigot (Univ. of Grenoble, CNRS) , Oudet (Univ. of Grenoble)**q

  A large class of non-linear continuity equations with confinement and/or possibly non local interaction potential can be considered as semi discrete gradient flows with respect to the Euclidean Wassertein distance. The numerical resolution of such problem in dimension 2 and higher is open. Our approach is based on two remarks : the reformulation of the optimization problem in terms of Brenier potential seems to behave better. This introduces a Monge-Ampère operator in the cost functional which needs a monotone discretization in order to preserve the convexity at the discrete level. The first numerical results are very encouraging.

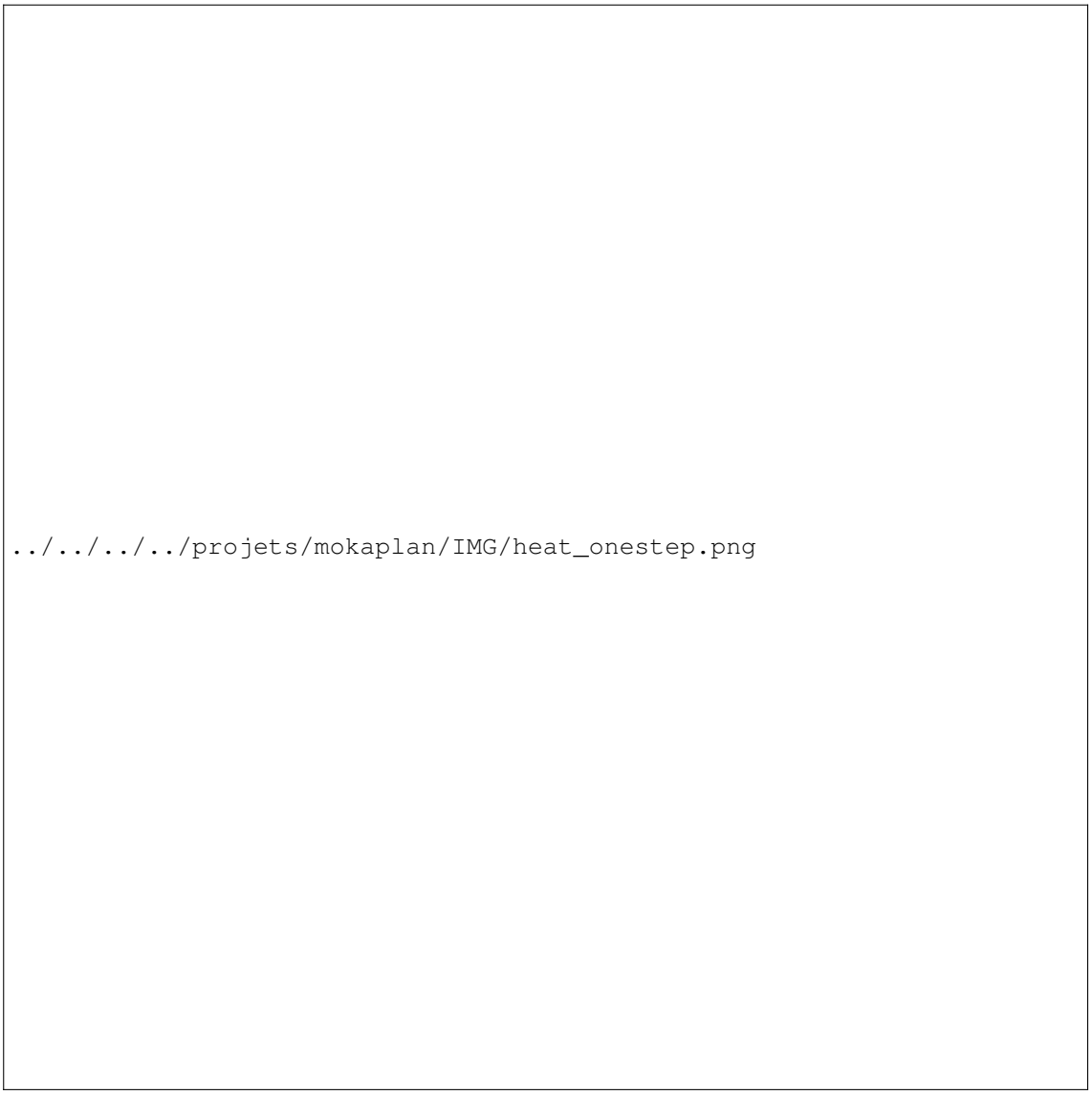


Figure 5. One step of Wasserstein JKO gradient flow for the classical entropy (our numerical method) compared to traditional Finite Difference of the heat equation. Left the initial heat profile, right the heat profile after one time step for both methods.

- **Benamou, Carlier, Agueh (Univ. of Victoria)** Splitting methods for kinetic equations, we try to use one JKO step to deal with the non-linear velocity advection part of kinetic equations [31]. This seems to be relevant to granular media equation [16], and also may offer a completely new method for Liouville equations arising from Geometrical Optics [19].

MUTANT Project-Team

6. New Results

6.1. Operational Timed Semantics

Participants: José Echeveste, Jean-Louis Giavitto, Florent Jacquemard, Arshia Cont.

One common use-case of real-time musical interactions between musicians and computers is *Automatic Accompaniment* where the system is comprised of a real-time machine listening system that in reaction to recognition of events in a score from a human performer, launches necessary actions for the accompaniment section. While the real-time detection of score events out of live musicians' performance has been widely addressed in the literature, score accompaniment (or the reactive part of the process) has been rarely discussed. In [13], we are trying to deal with this missing component in the literature from a programming language perspective. We show how language considerations would enable better authoring of time and interaction during programming/composing and how it addresses critical aspects of a musical performance (such as errors) in real-time. We sketch the real-time features required by automatic musical accompaniment seen as a reactive system and formalize the timing strategies for musical events taking into account the various temporal scales used in music. Various strategies for the handling of synchronization constraints and the handling of errors are presented.

The behavior of the system *Antescofo* have been formally modeled as a *network of parametric timed automata*. The model obtained provides operational semantics for the input scores, in particular the interaction between the instrumental and electronic parts and the timing and error handling strategies mentioned below. This approach enables better authoring of time and interaction during programing/composing, permitting to use state of the art software verification tools for the static analysis of Antescofo scores. It also provides means to address critical aspects of musical performances in real-time.

6.2. Timed Static Analysis of Interactive Music Scores

Participant: Florent Jacquemard.

It is well known that every musician performance of the same work will differ from another. It is therefore a challenging task to be able to predict the behavior of interactive music systems like *Antescofo* in response to any possible performance, and prevent unwanted outcomes. With Léa Fanchon, we have been working on a module for timing analysis of augmented scores that complements the real-time score authoring and performance in *Antescofo*, with the aim of exploring possible behavior of authored scores with respect to possible deviations in human musician performance.

For this purpose we have studied [24] the application of formal models and methods from the literature of real-time systems verification to the static analysis of interactive music systems. We have considered in particular the good parameters problem, which consists in synthesizing a set of timing parameter valuations (representing performances here) guarantying a good behavior of the system analyzed. The methods presented in [24] have been applied to *Antescofo*, providing the following input to users:

- Evaluation of robustness of the program with respect to the environment's (musician's performance) temporal variations,
- Feedback to programmers or artists on critical synchronization points for better programming.

This study is one of the first of this kind in computer music literature, and the methods presented are general enough to apply to the verification of other interactive multimedia applications.

6.3. Automating the Generation of Test Suites for Antescofo

Participants: Florent Jacquemard, Clément Poncelet.

Clément Poncelet has started to develop during his Master thesis [35] a framework for black box conformance testing of *Antescofo*. This work is pursued in a PhD supported by DGA and Inria. The most important task in this context is the generation of relevant test data for the system, given an augmented score in *Antescofo* language. This data includes input, containing musical events (notes, chords etc) together with their timings. In a sense, the input data simulates a musical execution of the score. The input data must then be passed to *Antescofo* for black-box execution, in order to observe the system's reactions and compare them the expected output. For the latter comparison task, we need to be able to define the expected output, hence to have a formal model of the expected behavior of the system on the given score. For this purpose, we are using models of the system made of timed automata, which are computed automatically from given music scores. Then, we use tools from the UPPAAL suite [40] in order to generate testing data, based on relevant covering criteria and a formal model of the environment (i.e. the musician). This work has been presented at the poster session of MSR 2013 (national colloquium on modeling reactive systems) and a journal paper is in preparation.

6.4. Synchronous Embedding of Antescofo DSL

Participants: Arshia Cont, Jean-Louis Giavitto, Florent Jacquemard.

Antescofo can been seen as the coupling of a listening machine and a real-time reactive system. Therefore, it faces some of the same major challenges as embedded systems. We have been working with Guillaume Baudart, Louis Mandel, and Marc Pouzet (EPI Parkas, ENS) in strengthening the ties between the reactive aspects of *Antescofo* and that of synchronous languages, in particular ReactiveML [44]. In [17], we present a synchronous semantics for the core language of Antescofo and an alternative implementation, based on an embedding inside the synchronous language ReactiveML [44]. The semantics reduces to a few rules, is mathematically precise and leads to an interpretor of a few hundred lines whose efficiency compares well with that of the current implementation. On all musical pieces we have tested, response times have been less than the reaction time of the human ear. Moreover, this embedding permitted the prototyping of several new programming constructs. Some examples are available, together with the ReactiveML source code at http://reactiveml.org/emsoft13/.

6.5. Tree Structured Representation of Symbolic Temporal Data

Participant: Florent Jacquemard.

In traditional music notation, in particular in the languages used for the notation of mixed music such as Antescofo DSL, the durations are not expressed by numerical quantities but by symbols representing successive subdivisions of a reference time value (the beat). For this reason, trees data structures are commonly used for the symbolic representation of rhythms in computer aided composition softwares such as OpenMusic (developed at Ircam).

Following this idea, we have been working on using several tree automata techniques for the challenging and long-standing problem of automatic transcription of rhythm (in traditional music notations) from symbolic input data (symbolic traces with timestamps in ms, like e.g. in MIDI format). To summarize, the main problem in rhythm transcription is to find an acceptable balance between timing precision (the goal is to minimize the loss obtained by transformation of ms timing values into fractions of beats) and the complexity of the notation obtained. The relative importance of these two measures may vary largely according to the user (composer), his workflow, and the musical style considered. It is therefore important to be able to control this balance during the transcription process, in order to adapt to the case of users. In traditional approaches, the transcription is done by an alignement of the input trace on a grid, and the two measures (precision of the grid and complexity) are either defined by parameters fixed a priori or hardcoded e.g. for a precise musical style and composition workflow. During two internships co-supervised by Jean Bresson (Ircam, main developer of OpenMusic) and Florent Jacquemard, we have been studying more flexible new approaches, based on computations on the tree representation of rhythms.

Pierre Donat-Bouillud (L3 ENS Rennes) [29] has worked on an approach by transformation of trees following some rewrite rules. The general idea is to start with a complex tree representing timings very close to the input data, and to simplify it by rewriting until an acceptable level of complexity is reached. The rewrite rules are either generic (defining an equational theory of rhythm notation) or user defined (defining approximations). This approach has been implemented in an OpenMusic library.

Adrien Maire (M1 ENS Cachan) has studied another very promising approach based on stochastic tree automata learning in an interactive authoring scenario. The generated automaton is supposed to represent (by the weighted tree langage it defines) the expected complexity of rhythm notations (i.e. the user's "style").

Moreover, we have following other work on several classes of tree recognizers and tree transformations which could be of interest in this context. With Luis Barguñó, Carlos Creus, Guillem Godoy, and Camille Vacher, [11] we define a class of ranked tree automata called TABG generalizing both the tree automata with local brother tests of Bogaert and Tison [37] and with global equality and disequality constraints (TAGED) of Filiot et al. [39]. TABG can test for equality and disequality modulo a given flat equational theory between brother subterms and between subterms whose positions are defined by the states reached during a computation. In particular, TABG can check that all the subterms reaching a given state are distinct. This constraint is related to monadic key constraints for XML documents, meaning that every two distinct positions of a given type have different values. We have proven decidability of the emptiness problem for TABG. This solves, in particular, the open question of decidability of emptiness for TAGED. We further extended our result by allowing global arithmetic constraints for counting the number of occurrences of some state or the number of different equivalence classes of subterms (modulo a given flat equational theory) reaching some state during a computation. We also adapt the model to unranked ordered terms. As a consequence of our results for TABG, we prove the decidability of a fragment of the monadic second order logic on trees extended with predicates for equality and disequality between subtrees, and cardinality.

With Michaël Rusinowitch (EPI Cassis), we have introduced in [25] an extension of unranked tree automata called bi-dimensional context-free hedge automata. The languages they define are context free in two dimensions: in the the sequence of successors of a node and also along paths. This formalism is useful for the static type-checking of tree transformations such as XML updates defined in the W3C XQuery Update Facility. We have developed with the same author in the past years a general framework for the verification of unranked (XML) tree transformations based on tree automata techniques. It has been presented this year in an invited keynote [16]. We have also presented with Emmanuel Filiot and Sophie Tison a survey on tree automata with constraints [33] during a Dagstuhl Seminar (number 13192) on tree transducers and formal methods.

6.6. Online Automatic Structure Discovery of Audio Signals

Participants: Arshia Cont, Vincent Lostanlen [MS Internship].

Following recent team findings in [12] and the framework introduced in [4], we pursued the problem of automatic discovery of audio signals using methods of information geometry through a Masters Thesis undertaken by Vincent Lostanlen (MS ATIAM) [34]. This work introduces a novel way of representing and calculating *Similarity Matrices* for continuous multimedia signals and in real-time. In this approach, the signal is first segmented into homogeneous chunks using the change detection algorithm proposed by the team in [12], and proposes a method for constituting similarity relations between segments using *Bregman Information Geometry* and exploiting intersections between information balls.

Compared to traditional approaches to similarity matrix computing, the approach proposed in [34] is strictly on-line (thus suitable for real-time computing) and provides a sparse view of audio structures. We will pursue this project by increasing its robustness and evaluating results on larger databases including other timed-signals such as video.

6.7. Temporal Coherency Criterion for Alignment Inference Algorithms

Participants: Philippe Cuvillier [PhD Student], Arshia Cont.

The question of modeling time and duration is of utmost importance for stability and robustness of real-time alignment algorithms and constitute one of the major success factors for the *Antescofo* listening machine described in [2]. Meanwhile, regular algorithms undergo stability in highly uncertain environments where observations obtained from the signal are highly uninformative and temporal information is of crucial importance.

PhD student Philippe Cuvillier defined *Coherency Criteria* for such applications and attempted to formalize such criteria in terms of probabilistic models and inference algorithms in case of Hidden Semi-Markov Chains. The results show that not all probabilistic families meet such criteria including some commonly used by engineers and designers. Preliminary results are submitted for publications and experimental results are being pursued.

PARKAS Project-Team

6. New Results

6.1. Reactive Programming

Participants: Guillaume Baudart, Louis Mandel, Cédric Pasteur, Marc Pouzet.

ReactiveML is an extension of OCaml with synchronous concurrency, based on synchronous parallel composition and broadcast of signals. The goal is to provide a general model of deterministic concurrency inside a general purpose functional language to program reactive systems. It is particularly suited to program discrete simulations, for instance of sensor networks.

One of the current focus of the research is being able to simulate huge systems, composed of millions of agents, by extending the current purely sequential implementation in order to be able to take advantage of multi-core and distributed architectures. This goal has led to the introduction of a new programming construct, *reactive domain*, which allows to define local time scales. These domains help for the distribution of the code but also increase the expressiveness of the language. In particular, it allows to do time refinement. A paper on this new construct and the related static analysis has been published [20]. An extended version is under submission.

We continued the work on a new reactivity analysis which ensures that a process can not prevent the other ones to from executing. This analysis has published in [19]. An English version is under submission.

The runtime of ReactiveML has been cleanup and a multi-threaded implementation has been developed. A paper describing this new implementation will be published in [27].

All these novelties has been described precisely in the PhD thesis of Cédric Pasteur [1].

During the year, ReactiveML has also bee applied to *mixed music*. Mixed music is about live musicians interacting with electronic parts which are controlled by a computer during the performance. It allows composers to use and combine traditional instruments with complex synthesized sounds and other electronic devices. There are several languages dedicated to the writing of mixed music scores. Among them, the Antescofo language coupled with an advanced score follower allows a composer to manage the reactive aspects of musical performances: how electronic parts interact with a musician. However these domain specific languages do not offer the expressiveness of functional programming.

We defined a synchronous semantics for the core language of Antescofo and an alternative implementation based on an embedding inside ReactiveML [9]. The semantics reduces to a few rules, is mathematically precise and leads to an interpretor of only a few hundred lines. The efficiency of this interpretor compares well with that of the actual implementation: on all musical pieces we have tested, response times have been less than the reaction time of the human ear. Moreover, this approach offers to the composer recursion, higher order, inductive types, as well as a simple way to program complex reactive behaviors thanks to the synchronous model of concurrency on which ReactiveML is built [10].

6.2. n-Synchronous Languages

Participants: Albert Cohen, Adrien Guatto, Louis Mandel, Marc Pouzet.

Synchronous programming languages in the vein of Lustre were designed for critical real-time systems. They are, however, not that well adapted to embedded applications with more pressing computational needs, since the generated code will usually not contain loops or arrays.

An essential task of a Lustre compiler is to determine whether a program can be executed within bounded memory. This process is called the "clock calculus", and consists in mapping every item of each program stream to a logical date in a global, discrete time scale. For a given stream, the mapping itself is called a "clock", and is a strictly increasing function from stream positions to natural numbers representing ticks: two items cannot be computed at the same time. In practice, this function is represented as an infinite binary stream where the boolean b_i denotes presence (or absence) in the corresponding data stream at the i-th instant.

In recent work, Guatto, Cohen, Mandel and Pouzet considered the extension of the Lustre and Lucid Synchrone clock calculus to allow computing several values instantaneously. This simple idea has a deep impact on all aspects of the language: - its denotational semantics has to account for bursts of values; - the clock calculus now features integers rather than booleans: each integer denotes the size of the burst at the corresponding instant; - causality analysis has to take bursts into account when rejecting self-referential programs; - the code generation process translates bursts to arrays and clocks to counted loops.

A prototype implementation exploiting this idea and generating C code with loops is underway and a paper describing the base of the clock calculus will be published [26].

This work extends nicely the n-synchronous model that introduced a way to compose streams which have *almost the same clock* and can be synchronized through the use of a finite buffer.

6.3. Mechanization of AODV loop freedom proof

Participant: Timothy Bourke.

The Ad hoc On demand Distance Vector (AODV) routing protocol is described in RFC3561. It allows the nodes in a Mobile Ad hoc Network (MANET) to know where to forward messages so that they eventually reach their destinations. The nodes of such networks are *reactive systems* that cooperate to provide a global service (the sending of messages from node to node) satisfying certain correctness properties (namely 'loop freedom'—that messages are never sent in circles).

We have mechanized an existing formal but pen-and-paper proof of loop freedom of AODV in the interactive theorem prover Isabelle/HOL. While the process algebra model and the fine details of the original proof are quite formal, the structure of the proof is much less so. This necessitated the development of new framework elements and techniques in Isabelle. In particular, we adapted standard theory on inductive assertions to show invariants over individual reactive nodes and introduced machinery for assume/guarantee reasoning to lift these invariants to networks of communicating processes. While the original proof reasoned informally over traces, the mechanized proof is purely based on invariant reasoning, i.e., on reasoning over pairs of reachable states. Our combination of techniques works very well and is likely useful for modelling and verifying similar protocols in an interactive theorem prover.

We are currently finalising a paper describing this work for submission in January.

In collaboration with Peter Hofner (NICTA) and Robert J. van Glabbeek (UNSW/NICTA).

6.4. Hybrid Synchronous Languages

Participants: Timothy Bourke, Jun Inoue, Antoine Madet, Marc Pouzet.

During year 2013, we mainly worked on three directions: (a) the treatment of DAEs; (b) the design and implementation of a causality analysis for hybrid systems modelers; (c) the study of numerical techniques for *non-smooth dynamical systems*.

DAEs As part of our participation in the European project MODRIO and SYS2SOFT projects, we have been developing a prototype for simulating DAE (Differential-Algebraic Equations) systems. DAEs are the basis of the language Modelica and their interaction with discrete features — in particular the novel ones introduced in 2012, like hierarchical automata and clocks — raise difficult semantical and compilation issues. The goal is to precisely define the interaction between synchronous programming constructs and DAEs, in term of semantics and compilation. One strong difficulty at the moment is that existing techniques (index reduction, dymmy derivative) are not modular and force, either to (a)

write an interpretor where index reduction is done dynamically every time a mode change occurs or (b) statically enumerate all the modes, performing index reduction for every of those. While the first technique is too slow in practice (and it is not used in the most advanced Modelica compiler), the second one may explode in practice (putting n two-state automata in parallel lead to 2^n states to be enumerated). During year 2013, we have investigated a new approach for index reduction.

Work to-date has focused on implementing standard algorithms from the literature (notably Pantelides, Dummy Derivatives, Dynamic State Selection). Despite the importance of these algorithms to tools like Modelica, we found that important implementation details and "tricks" are not always well documented.

This work is developed hand-in-hand with the interface to the Sundials IDA solver.

Causality Analysis We have designed a causality analysis for a language that mix stream equations, hierarchical automata and ODEs and implemented it in the Zélus compiler. Its purpose is to give a sufficient condition for a hybrid program can be turned into statically scheduled code. Moreover, the analysis ensures that absence of discontinuities outside of declared zero-crossing events. This result is novel and the proof deeply rely on the use of *non standard analysis* introduced in our previous works. This new result has been accepted for publication at HSCC 2014.

Non Smooth Dynamical Systems In parallel, we collaborate with Bernard Brogliato and Vincent Acary (Inria team BIBOP, Grenoble) on non smooth dynamical systems. Beside general-purpose techniques for solving DAEs and implemented in Modelica compilers, there exist dedicated methods for systems with a lot of discontinuities and contacts (in mechanical system, electrical analogous circuits, etc.). They are far more efficient and numerically accurate than general-purpose techniques when the number of contact is important (e.g., transient in electrical circuits, a bag of marbles). They are based on a time stepping execution and do not have to stop at every zero-crossing event. The combination of those techniques with event detection ones (as used in the Simulink tool) is largely unknown. We are currently inverstigating the extension of our previous work to take Brogliato and Acary techniques into account. This is a novel but promising direction of research for the year to come.

In this research activity, we develop the new language Zélus used as a laboratory for experimenting novel programming constructs and compilation techniques. It serves to illustrate our research as Lucid Synchrone did in the past.

In collaboration with Benoit Caillaud and Albert Benveniste of the Inria HYCOMES team.

6.5. Fidelity in Real-Time Programming

Participants: Timothy Bourke, Guillaume Baudart.

We are close to completing a careful analysis of literature related to the quasi-synchronous model for real-time, distributed systems. We have extended existing results by increasing their precision, providing detailed proofs, and simplifying protocol descriptions. The work to-date is documented in a draft document which we expect will eventually become a technical report or journal article.

Quasi-synchronous architectures, sometimes termed Loosely Time-Triggered Architectures (LTTAs), are ubiquitious in the development of distributed, real-time systems. They represent a broad class of systems whose modelling and programming mixes elements of discrete time, physical time, and a notion of approximation. We expect that addressing these elements—in the Zélus programming language—will lead to insights and advances in a broader ambition to program in physical time.

6.6. A theory of safe optimisations in the C11/C++11 memory model and applications to compiler testing

Participants: Francesco Zappa Nardelli, Robin Morisset.

Compilers sometimes generate correct sequential code but break the concurrency memory model of the programming language: these subtle compiler bugs are observable only when the miscompiled functions interact with concurrent contexts, making them particularly hard to detect. In this work we design a strategy to reduce the hard problem of hunting concurrency compiler bugs to differential testing of sequential code and build a tool that puts this strategy to work. Our first contribution is a theory of sound optimisations in the C11/C++11 memory model, covering most of the optimisations we have observed in real compilers and validating the claim that common compiler optisations are sound in the C11/C++11 memory model. Our second contribution is to show how, building on this theory, concurrency compiler bugs can be identified by comparing the memory trace of compiled code against a reference memory trace for the source code. Our tool identified several mistaken write introductions and other unexpected behaviours in the latest release of the gcc compiler.

A paper on this work has been accepted in [22].

6.7. A verified compiler for relaxed-memory concurrency

Participant: Francesco Zappa Nardelli.

We studied the semantic design and verified compilation of a C-like programming language for concurrent shared-memory computation above x86 multiprocessors. The design of such a language is made surprisingly subtle by several factors: the relaxed-memory behaviour of the hardware, the effects of compiler optimisation on concurrent code, the need to support high-performance concurrent algorithms, and the desire for a reasonably simple programming model. In turn, this complexity makes verified (or verifying) compilation both essential and challenging. This project started in 2010. In 2013 an article, describing the correctness proof of all the phases of our CompCertTSO compiler (including experimental fence eliminations), appeared in the Journal of the ACM [7].

In collaboration with Jaroslav Sevcik (U. Cambridge), Viktor Vafeiadis (MPI-SWS), Suresh Jagannathan (Purdue U.), Peter Sewell (U. Cambridge).

6.8. Language design on top of JavaScript

Participant: Francesco Zappa Nardelli.

This research project aims at improving the design of the JavaScript language. In [23] we present a security infrastructure which allows users and content providers to specify access control policies over subsets of a JavaScript program by leveraging the con- cept of delimited histories with revocation. We implement our proposal in WebKit and evaluate it with three policies on 50 widely used websites with no changes to their JavaScript code and report performance overheads and violations. In [32] 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 and we report on performance and software engineering benefits.

With Gregor Richards and Jan Vitek (Purdue University).

6.9. Tiling for iterated stencils

Participants: Tobias Grosser, Sven Verdoolaege, Albert Cohen.

Time-tiling is necessary for the efficient execution of iterative stencil computations. Classical hyper-rectangular tiles cannot be used due to the combination of backward and forward dependences along space dimensions. Existing techniques trade temporal data reuse for inefficiencies in other areas, such as load imbalance, redundant computations, or increased control flow overhead, therefore making it challenging for use with GPUs.

We proposed a time-tiling method for iterative stencil computations on GPUs. Our method is the first tiling algorithm solving the following constraints simultaneously: it does not involve redundant computations, it favors coalesced global-memory accesses, data reuse in local/shared-memory or cache, avoidance of thread divergence, and concurrency, combining hexagonal tile shapes along the time and one spatial dimension with classical tiling along the other spatial dimensions. Hexagonal tiles expose multi-level parallelism as well as data reuse. Experimental results demonstrate significant performance improvements over existing stencil compilers.

Part of this work also involved our colleagues from the POLYFLOW associate-team at the Indian Institute of Science, Bangalore, India.

6.10. Compilation for scalable on-chip parallelism

Participants: Antoniu Pop, Feng Li, Sven Verdoolaege, Govindarajan Ramaswamy, Albert Cohen.

Task-parallel programming models are getting increasingly popular. Many of them provide expressive mechanisms for inter-task synchronization. For example, OpenMP 4.0 will integrate data-driven execution semantics derived from the StarSs research language. Compared to data-parallel and fork-join models of parallelism, the advanced features being introduced into task-parallel models in turn enable improved scalability through load balancing, memory latency mitigation, mitigation of the pressure on memory bandwidth, and as a side effect, reduced power consumption.

We developed a systematic approach to compile a loop nest into concurrent, dependent tasks. We formulated a partitioning scheme based on the tile-to-tile dependences, represented as affine polyhedra. This scheme ensures at compilation time that tasks belonging to the same class have the same, fully explicit incoming and outgoing dependence patterns. This alleviates the burden of a full-blown dependence resolver to track the readiness of tasks at run time. We evaluated our approach and algorithms in the PPCG compiler, targeting OpenStream, our experimental data-flow task-parallel language with explicit inter-task dependences and a lightweight runtime. Experimental results demonstrate the effectiveness of the approach.

Part of this work also involved our colleagues from the POLYFLOW associate-team at the Indian Institute of Science, Bangalore, India.

6.11. Correct and efficient runtime systems

Participants: Nhat Minh Lê, Robin Morisset, Adrien Guatto, Antoniu Pop, Francesco Zappa Nardelli, Albert Cohen.

User-space scheduling and concurrent first-in first-out queues are two essential building blocks of parallel programming runtimes. They are, however, rarely used together since typical schedulers are oblivious to the ordering constraints introduced by buffered communication.

Chase and Lev's concurrent deque is a key data structure in shared-memory parallel programming and plays an essential role in work-stealing schedulers. We provided the first correctness proof of an optimized implementation of Chase and Lev's deque on top of the POWER and ARM architectures: these provide very relaxed memory models, which we exploit to improve performance but considerably complicate the reasoning. We also studied an optimized x86 and a portable C11 implementation, conducting systematic experiments to evaluate the impact of memory barrier optimizations. Our results demonstrate the benefits of hand tuning the deque code when running on top of relaxed memory models.

Based on this early success, we started working on a more global solution using a new lock-free algorithm for stalling and waking-up tasks in a user-space scheduler according to changes in the state of the corresponding queues. The algorithm is portable and correct, since it is written and proven against the C11 memory model. We showed through experiments that it can serve as a keystone to efficient parallel runtime systems.

These efforts underline the parallelizing compilation research for n-synchronous languages, and the scalable parallel execution of OpenStream.

6.12. Checking Synchronous Compiler Correctness

Participants: Francesco Zappa Nardelli, Guillaume Chelfi, Marc Pouzet.

During year 2013, we have worked on the use of formal verification of compilation steps in the compiler of a Lustre-like synchronous language. Two main directions has been taken:

- The use of SMT-based k-induction techniques to verify the correctness of the successive steps of a synchronous compiler. We used the tool KIND developed by Cesare Tinelli (Iowa state Univ.) and applied it to the Heptagon compiler. The compiler does several source-to-source transformations upto sequential code and KIND was used to verify the equivalence between those successive steps. We came to the conclusion that for most programs, equivalence checking fails unless extra traceability information is added by the compiler.
- The development of a dedicated verification technique to prove the equivalence between a Lustre program and its sequential implementation. We plan to pursue this work during year 2014. Cesare Tinelli will be visiting professor for a month during June 2014.

PI.R2 Project-Team

5. New Results

5.1. Proof-theoretical and effectful investigations

Participants: Pierre Boutillier, Guillaume Claret, Pierre-Louis Curien, Yann-Régis Gianas, Hugo Herbelin, Guillaume Munch-Maccagnoni, Ludovic Patey, Pierre-Marie Pédrot, Alexis Saurin.

5.1.1. Sequent calculus and computational duality

Categorical semantics.

During his collaboration with Marcelo Fiore and Pierre-Louis Curien, Guillaume Munch-Maccagnoni characterised the polarised evaluation order through a categorical structure where the hypothesis that composition is associative is relaxed. Duploid is the name of the structure, as a reference to Jean-Louis Loday's duplicial algebras. The main result, in the lineage of Führmann's [38] direct-style characterisation of monadic models, is a reflection Adj→Dupl where Dupl is a category of duploids and duploid functors, and Adj is the category of adjunctions and pseudo maps of adjunctions. The result suggests that the various biases in denotational semantics: indirect, call-by-value, call-by-name... are a way of hiding the fact that composition is not always associative. This work was accepted for publication in FoSSaCs 2014 [53].

Pierre-Louis Curien, in connection with his increasing interests in operads and algebraic structures of various kinds, found out that the core syntax of system L (underlying the duality of computation) could be used with profit to describe the wiring structures underlying operads, dioperads, cyclic operads, and more generally Lamarche's structads [48]. He also showed a syntactic equivalence between Munch-Maccagnoni's (pre)duploids and system L syntax. These results were presented in his invited talks at the Loday's Mathematical Legacy workshop in Strasbourg and at the workshop Algebra and Computation in Lyon, in January 2014.

Duality of construction.

Paul Downen and Zena Ariola developed a generalized theory of the sequent calculus for understanding the concepts of evaluation strategy and of data (for example, pairs in ML) and co-data (for example, functions) in programming languages. This theory provides a single framework for user-defined data and co-data types as well as a generalized treatment of evaluation strategies, including call-by-value, call-by-name, and call-by-need, that are given as parameters to the theory. In the end, the framework encompasses the previously known duality of call-by-name and call-by-value in the sequent calculus, both by Curien and Herbelin [3] and Wadler [59] , while also including call-by-need and its dual. Additionally, the framework reveals connections with approaches by Zeilberger [60], Munch-Maccagnoni [6], and Curien and Munch-Maccagnoni [33], for using polarization and focalization to provide deterministic strategies for classical computation with structures and pattern matching. This work will be presented at ESOP 2014 [15].

Luke Maurer and Zena Ariola in collaboration with Daniele Varacca studied the connections between π-calculus encodings of the λ-calculus and similar continuation-passing style (CPS) transformations, extending the connections for call-by-value and call-by-name encodings to include the call-by-need π-calculus encoding as well. This development revealed a better understanding of the computational effect needed in the λ-calculus to model call-by-need evaluation, which better reflects the way that memoization for call-by-need is implemented. The work is going to be submitted to RTA-TLCA.

Constructive interpretation of an involutive negation.

Guillaume Munch-Maccagnoni developped a syntax of delimited control operators that exposes a formulae-as-types correspondence between an involutive negation in classical natural deduction, and the idea that captured contexts, unlike continuations, can be inspected. This decomposes technical artefacts found in call-by-name classical realisability, and simplifies witness extraction from proofs of Σ formulae. This work has been submitted and appears in his PhD thesis [11].

5.1.2. Dependent monads

Guillaume Claret and Yann Régis-Gianas are developping a monadic translation from functional programs with effects to Coq that uses a dependent monad. The aim of this work is to allow to reason about effectful programs directly in Coq.

5.1.3. Linear dependent types

Pierre-Marie Pédrot developped a dependent version of the Dialectica translation, that gives interesting insights into the possibility to design linear dependent types. Indeed, Dialectica can be decomposed as a translation acting on linear types instead of intuitionistic ones.

5.1.4. Delimited continuations, polarity and computational effects

Guillaume Munch-Maccagnoni's polarised decomposition of delimited control calculi appeared in his PhD thesis [11].

5.1.5. Reverse mathematics

Ludovic Patey studied with Laurent Bienvenu and Paul Shafer the deep connections between algorithmic randomness and reverse mathematics by defining formally the ability of computing a solution to a problem by probabilistic means within the framework of reverse mathematics, the No Randomized Algorithm property (NRA). They provided a classification of the whole revese mathematics zoo created by Damir Dzhafarov in terms of having the NRA property or not, answering to some open separation questions.

Ludovic Patey stated two dichotomy theorems about satisfiability problems within reverse mathematics and proved them using clone theory. The corresponding paper is submitted to Computability in Europe 2014. He studied also ramseyan theorems related to the Rainbow Ramsey theorem and provided characterizations in terms of diagonally non-computable functions, algorithmic randomness, and related it to the Erdös Moser theorem and Thin set theorem.

5.1.6. Gödel's functional interpretation

Pierre-Marie Pédrot showed that the Dialectica translation could be explained in terms of the Krivine abstract machine, in a way similar to the usual presentation of classical realizability. This opens the door to a better understanding of related translations, as well as adding semi-classical effect into PTS.

5.1.7. Logical foundations of call-by-need evaluation

Alexis Saurin and Pierre-Marie Pédrot developed a structured reconstruction of call-by-need based on linear head reduction which arose in the context of linear logic. This opens new directions both to extend call-by-need to control and to apply linear logic proof-theory (and particularly proof-nets) to call-by-need evaluation.

5.1.8. Streams and classical logic

Alexis Saurin and Fanny He have been working on transfinite term rewriting in order to model stream calculi and their connections with lambda-calculi for classical logic.

Jaime Gaspar identified the eight simplest variants (some already known) of the Kuroda negative translation that translate classical logic into minimal logic.

5.2. Type theory and the foundations of Coq

Participants: Pierre Boutillier, Pierre-Louis Curien, Hugo Herbelin, Pierre-Marie Pédrot, Yann Régis-Gianas, Alexis Saurin, Matthieu Sozeau.

5.2.1. Substitutions and isomorphisms

Pierre-Louis Curien completed his joint work with Richard Garner and Martin Hofmann on relating syntax unstrictification through coercions with model strictification (cf. πr^2 report 2012), adding a careful treatment of identity types. The corresponding paper was accepted for publication in the TCS special issue for Glynn Winskel's anniversary.

5.2.2. Homotopy type theory

Hugo Herbelin, Matthieu Sozeau and Pierre-Louis Curien participated to the univalent foundations program. A collaborative book [18] on the results of this program has been published.

5.2.3. Models of type theory

Simplicial sets and their extensions as Kan complexes can serve as models of homotopy type theory. Hugo Herbelin developed a concrete type-theoretic formalisation of semi-simplicial sets following ideas from Steve Awodey, Peter LeFanu Lumsdaine and other researchers both at Carnegie-Mellon University and at the Institute of Advanced Study. This has been accepted for publication in a special issue of MSCS on homotopy type theory [22].

The technique he used generalises to provide type-theoretic constructions for arbitrary presheaves on Reedy categories, thus including simplicial sets. In particular, this provides with a formulation of simplicial sets where degeneracies are decidable, which is not the case with the definition as a presheaf.

Hugo Herbelin also investigated hybrid constructive definitions of simplicial sets where face maps are axiomatised but degeneracies are built. Again, this provides with a formulation of simplicial sets where it is decidable whether a given simplex is degenerate or not.

5.2.4. Internalizing the setoid model of type theory

As an example use of the new polymorphic universe extension of Coq, Matthieu Sozeau developed together with Nicolas Tabareau (Inria Ascola team, École des mines Nantes) a complete groupoid model of type theory, following the seminal work of Hofmann and Streicher. A preliminary paper presenting a partial generalization of this model to 2-goupoids was written and will be resubmitted [23].

A completed version of this model has since been formalized and will be submitted shortly. This model showcases the use of the polymorphic universes: in the course of its formalization we uncovered hidden assumptions in the interpretation of substitution and sigma types in the original presentation thanks to the universe system.

5.2.5. Proof irrelevance, eta-rules

Matthieu Sozeau finished his implementation of a proof-irrelevant system but did not publish it. Indeed, the homotopy type theory interpretation suggests new ways to introduce proof-irrelevance using bracket types that seem to significantly depart from the syntactic treatment developped by Werner and himself. An investigation of the relationship between the presentation of the calculus of inductive constructions given by Hugo Herbelin and Arnaud Spiwack in [44] which includes the bracket construction and the aforementioned syntactic version will be part of a master's internship supervised by Matthieu Sozeau in 2014.

5.3. Homotopy of rewriting systems

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos.

5.3.1. The homotopical completion-reduction procedure

In [39], Stéphane Gaussent (Institut Camille Jordan), Yves Guiraud and Philippe Malbos have introduced the homotopical completion-reduction procedure as a higher-dimensional rewriting method to compute coherent presentations of monoids. The results of this procedure on Artin monoids of spherical type have been implemented by Yves Guiraud in a Python library, available on his webpage. The procedure is currently improved towards the explicit computation of full polygraphic resolutions of Artin monoids to provide a purely algebraic and constructive account of well-known geometric objects, such as Caylay graphs and Salvetti complexes.

In [16], Yves Guiraud, Philippe Malbos and Samuel Mimram (CEA Saclay) have further investigated the homotopical completion-reduction procedure, extended with the adjunction/elimination of redundant generators, with successful application to two new classes of monoids: the plactic and the Chinese monoids. This work has been implemented by Samuel Mimram and Yves Guiraud into a prototype, that can be tested at http://www.pps.univ-paris-diderot.fr/~smimram/rewr, and has been presented to RTA 2013 by Philippe Malbos, where it has received the best paper award.

5.3.2. New methods for the computation of coherent presentations

During his M2 internship, Maxime Lucas, supervised by Yves Guiraud, has improved the rewriting method used in [43] for the computation of homotopy bases of monoids and categories. This allows a more effective computation in several cases, based on the notion of Anick chain [25] instead of the broader notion of critical branching. Maxime Lucas has now started a PhD thesis, supervised by Yves Guiraud and Pierre-Louis Curien, and currently investigates the use of Garside-like structures [35] to further improve the computation of coherent presentations for higher-dimensional categories.

5.3.3. Higher-dimensional linear rewriting

Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud and Philippe Malbos investigate with Eric Hoffbeck (LAGA, Université Paris 13) and Samuel Mimram (CEA Saclay) the links between set-theoretic rewriting theory and the computational methods known in symbolic algebra, such as Gröbner bases [28]. This interaction is supported by the Focal project of the IDEX Sorbonne Paris Cité. Yves Guiraud, Eric Hoffbeck and Philippe Malbos are currently working on an improvement, based on the homotopical completion-reduction procedure, of the methods known in algebra to compute homological invariants of algebras and operads. Cyrille Chenavier has started a PhD thesis, supervised by Yves Guiraud and Philippe Malbos, to use Berger's theory of reduction operators [27] to design new methods for the study of rewriting systems.

5.3.4. Homotopical and homological finiteness conditions

Yves Guiraud and Philippe Malbos have written a comprehensive introduction [21] on the links between higher-dimensional rewriting, the homotopical finiteness condition "finite derivation type" and the homological finiteness condition "FP_3", from the point of view of higher categories and polygraphs. The purpose of this work is to provide an introduction to the field, formulated in a contemporary language, and with new, more formal proofs of classical results.

In [19], Yves Guiraud and Philippe Malbos have introduced a notion of identities among relations for higher categories presented by polygraphs. This notion is well-known in combinatorial group theory, where it is linked to the explicit computation of homological invariants and of formal representations of groups as crossed complexes. The main result of [19] is a procedure based on higher rewriting to compute generators of the identities among relations. They have related the facts that the natural system of identities among relations is finitely generated and that the higher category has finite derivation type (a homotopical finiteness condition introduced in [43] for higher categories after Squier's work for monoids [57]).

5.4. Coq as a functional programming language

Participants: Pierre Boutillier, Guillaume Claret, Lourdes Del Carmen González Huesca, Hugo Herbelin, Pierre Letouzey, Matthias Puech, Yann Régis-Gianas, Matthieu Sozeau.

5.4.1. Type classes and libraries

Type Classes are heavily used in the HoTT/Coq library (http://github.com/HoTT/coq) developed by the Univalent Foundations program at the IAS, to which Matthieu Sozeau participated.

5.4.2. Dependent pattern-matching

How to encode structurally dependent pattern matching into case analysis by hand has been described by Jean François Monin in [52]. Pierre Boutillier, with the help of Thomas Braibant (GALLIUM team), has mechanized this process and exhibited a missing part to make it scale. These are the main results presented in Pierre Boutillier's forthcoming thesis.

5.4.3. Incrementality in proof languages

Lourdes González and Yann Régis-Gianas studied incremental computing and self-adjusting computation [24] as a starting point to develop an applicative notion of change over data structures, to be applied to lambda-terms. They formalized in Coq a notion of derivative of an inductive function to define how to compute a new result from an input that has changed, this is done by using the derivative of the function and the difference on inputs and old outputs. They are working out a technique that allows a specification of functions using derivatives and old inputs and outputs including a cost analysis of the benefits of reusing previous computations.

5.4.4. Lightweight proof-by-reflection

In collaboration with Beta Ziliani (MPI), In the context of the ANR project Paral-ITP, Lourdes del Carmen González Huesca, Guillaume Claret and Yann Régis-Gianas developed a new technique for proof-by-reflection based on a notion of *a posteriori* simulation of effectful computations in Coq. This work has been presented at ITP 2013 ([14]).

POLSYS Project-Team

6. New Results

6.1. Fundamental Algorithms and Structured Systems

6.1.1. Structured polynomial systems: the quasi-homogeneous case

Let $\mathbb{K}$ be a field and $(f_1, ..., f_n) \subset \mathbb{K}[X_1, ..., X_n]$ be a sequence of quasi-homogeneous polynomials of respective weighted degrees $(d_1, ..., d_n)$ w.r.t a system of weights $(w_1, \cdots, w_n)$. Such systems are likely to arise from a lot of applications, including physics or cryptography. In [29], we design strategies for computing Gröbner bases for quasi-homogeneous systems by adapting existing algorithms for homogeneous systems to the quasi-homogeneous case. Overall, under genericity assumptions, we show that for a generic zero-dimensional quasi-homogeneous system, the complexity of the full strategy is polynomial in the weighted Bézout bound $\prod_{i=1}^{n} d_i / \prod_{i=1}^{n} w_i$. We provide some experimental results based on generic systems as well as systems arising from a cryptography problem. They show that taking advantage of the quasi-homogeneous structure of the systems allow us to solve systems that were out of reach otherwise.

6.1.2. Structured polynomial systems: the determinantal case

In [13], We study the complexity of solving the *generalized MinRank problem*, i.e. computing the set of points where the evaluation of a polynomial matrix has rank at most r. A natural algebraic representation of this problem gives rise to a *determinantal ideal*: the ideal generated by all minors of size $r + 1$ of the matrix. We give new complexity bounds for solving this problem using Gröbner bases algorithms under genericity assumptions on the input matrix. In particular, these complexity bounds allow us to identify families of generalized MinRank problems for which the arithmetic complexity of the solving process is polynomial in the number of solutions. We also provide an algorithm to compute a rational parametrization of the variety of a 0-dimensional and radical system of bi-degree $(D, 1)$. We show that its complexity can be bounded by using the complexity bounds for the generalized MinRank problem.

6.1.3. On the Complexity of the Generalized MinRank Problem

In [13] we study the complexity of solving the *generalized MinRank problem*, i.e. computing the set of points where the evaluation of a polynomial matrix has rank at most r. A natural algebraic representation of this problem gives rise to a *determinantal ideal*: the ideal generated by all minors of size $r + 1$ of the matrix. We give new complexity bounds for solving this problem using Gröbner bases algorithms under genericity assumptions on the input matrix. In particular, these complexity bounds allow us to identify families of generalized MinRank problems for which the arithmetic complexity of the solving process is polynomial in the number of solutions. We also provide an algorithm to compute a rational parametrization of the variety of a 0-dimensional and radical system of bi-degree $(D, 1)$. We show that its complexity can be bounded by using the complexity bounds for the generalized MinRank problem.

6.1.4. On the Complexity of Computing Gröbner Bases for Quasi-homogeneous Systems

Let $\mathbb{K}$ be a field and $(f_1, ..., f_n) \subset \mathbb{K}[X_1, ..., X_n]$ be a sequence of quasi-homogeneous polynomials of respective weighted degrees $(d_1, ..., d_n)$ w.r.t a system of weights $(w_1, \cdots, w_n)$. Such systems are likely to arise from a lot of applications, including physics or cryptography.

In [29], we design strategies for computing Gröbner bases for quasi-homogeneous systems by adapting existing algorithms for homogeneous systems to the quasi-homogeneous case. Overall, under genericity assumptions, we show that for a generic zero-dimensional quasi-homogeneous system, the complexity of the full strategy is polynomial in the weighted Bézout bound $\prod_{i=1}^{n} d_i / \prod_{i=1}^{n} w_i$.

We provide some experimental results based on generic systems as well as systems arising from a cryptography problem. They show that taking advantage of the quasi-homogeneous structure of the systems allow us to solve systems that were out of reach otherwise.

6.1.5. Gröbner bases of ideals invariant under a commutative group : the non-modular case

In [30], we propose efficient algorithms to compute the Gröbner basis of an ideal $I \subset k[x_1, \cdots, x_n]$ globally invariant under the action of a commutative matrix group G, in the non-modular case (where $char(k)$ doesn't divide $|G|$). The idea is to simultaneously diagonalize the matrices in G, and apply a linear change of variables on I corresponding to the base-change matrix of this diagonalization. We can now suppose that the matrices acting on I are diagonal. This action induces a grading on the ring $R = k[x_1, \cdots, x_n]$, compatible with the degree, indexed by a group related to G, that we call G-degree. The next step is the observation that this grading is maintained during a Gröbner basis computation or even a change of ordering, which allows us to split the Macaulay matrices into $|G|$ submatrices of roughly the same size. In the same way, we are able to split the canonical basis of R/I (the staircase) if I is a zero-dimensional ideal. Therefore, we derive *abelian* versions of the classical algorithms F_4, F_5 or FGLM. Moreover, this new variant of F_4/F_5 allows complete parallelization of the linear algebra steps, which has been successfully implemented. On instances coming from applications (NTRU crypto-system or the Cyclic-n problem), a speed-up of more than 400 can be obtained. For example, a Gröbner basis of the Cyclic-11 problem can be solved in less than 8 hours with this variant of F_4. Moreover, using this method, we can identify new classes of polynomial systems that can be solved in polynomial time.

6.1.6. Signature Rewriting in Gröbner Basis Computation

In [27] we introduce the RB algorithm for Gröbner basis computation, a simpler yet equivalent algorithm to F5GEN. RB contains the original unmodified F5 algorithm as a special case, so it is possible to study and understand F5 by considering the simpler RB. We present simple yet complete proofs of this fact and of F5's termination and correctness. RB is parametrized by a rewrite order and it contains many published algorithms as special cases, including SB. We prove that SB is the best possible instantiation of RB in the following sense. Let X be any instantiation of RB (such as F5). Then the S-pairs reduced by SB are always a subset of the S-pairs reduced by X and the basis computed by SB is always a subset of the basis computed by X.

6.1.7. An analysis of inhomogeneous signature-based Gröbner basis computations

In [8] we give an insight into the behaviour of signature-based Gröbner basis algorithms, like F5, G2V or SB, for inhomogeneous input. On the one hand, it seems that the restriction to sig-safe reductions puts a penalty on the performance. The lost connection between polynomial degree and signature degree can disallow lots of reductions and can lead to an overhead in the computations. On the other hand, the way critical pairs are sorted and corresponding s-polynomials are handled in signature- based algorithms is a very efficient one, strongly connected to sorting w.r.t. the well-known sugar degree of polynomials.

6.1.8. Improving incremental signature-based Gröbner basis algorithms

In [9] we describe a combination of ideas to improve incremental signature-based Gröbner basis algorithms having a big impact on their performance. Besides explaining how to combine already known optimizations to achieve more efficient algorithms, we show how to improve them even more. Although our idea has a positive affect on all kinds of incremental signature-based algorithms, the way this impact is achieved can be quite different. Based on the two best-known algorithms in this area, F5 and G2V, we explain our idea, both from a theoretical and a practical point of view.

6.1.9. A new algorithmic scheme for computing characteristic sets

Ritt-Wu's algorithm of characteristic sets is the most representative for triangularizing sets of multivariate polynomials. Pseudo-division is the main operation used in this algorithm. In [18] we present a new algorithmic scheme for computing generalized characteristic sets by introducing other admissible reductions than pseudo-division. A concrete subalgorithm is designed to triangularize polynomial sets using selected admissible reductions and several effective elimination strategies and to replace the algorithm of basic sets (used in Ritt-Wu's algorithm). The proposed algorithm has been implemented and experimental results show that it

performs better than Ritt-Wu's algorithm in terms of computing time and simplicity of output for a number of non-trivial test examples

6.2. Solving Systems over the Reals and Applications

6.2.1. On the Boolean complexity of real root refinement

In [32] 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 $\widetilde{O}_B(d^2\tau + dL)$. Our algorithms support the same complexity bound for the refinement of r roots, for any $r \leq d$.

6.2.2. On the minimum of a polynomial function on a basic closed semialgebraic set and applications

In [17] we give an explicit upper bound for the algebraic degree and an explicit lower bound for the absolute value of the minimum of a polynomial function on a compact connected component of a basic closed semialgebraic set when this minimum is not zero. We also present extensions of these results to non-compact situations. As an application, we obtain a lower bound for the separation of two disjoint connected components of basic closed semialgebraic sets, when at least one of them is compact.

6.2.3. Rational solutions to Linear Matrix Inequalities and Sums of Squares

Consider a $(D \times D)$ symmetric matrix A whose entries are linear forms in $\mathbb{Q}[X_1, ..., X_k]$ with coefficients of bit size $\leq \tau$. In [31], we provide an algorithm which decides the existence of rational solutions to the linear matrix inequality $\mathsf{A} \succeq 0$ and outputs such a rational solution if it exists. This problem is of first importance: it can be used to compute algebraic certificates of positivity for multivariate polynomials. Our algorithm runs within $(k\tau)^{O(1)}2^{O(\min(k,D)D^2)}D^{O(D^2)}$ bit operations; the bit size of the output solution is dominated by $\tau^{O(1)}2^{O(\min(k,D)D^2)}$. These results are obtained by designing algorithmic variants of constructions introduced by Klep and Schweighofer. This leads to the best complexity bounds for deciding the existence of sums of squares with rational coefficients of a given polynomial. We have implemented the algorithm; it has been able to tackle Scheiderer's example of a multivariate polynomial that is a sum of squares over the reals but not over the rationals; providing the first computer validation of this counter-example to Sturmfels' conjecture.

6.2.4. Exact Voronoi diagram of smooth convex pseudo-circles: General predicates, and implementation for ellipses

In [10] we examine the problem of computing exactly the Voronoi diagram (via the dual Delaunay graph) of a set of, possibly intersecting, smooth convex pseudo-cirlces in the Euclidean plane, given in parametric form. Pseudo-circles are (convex) sites, every pair of which has at most two intersecting points. The Voronoi diagram is constructed incrementally. Our first contribution is to propose robust and efficient algorithms, under the exact computation paradigm, for all required predicates, thus generalizing earlier algorithms for non-intersecting ellipses. Second, we focus on INCIRCLE, which is the hardest predicate, and express it by a simple sparse 5×5 polynomial system, which allows for an efficient implementation by means of successive Sylvester resultants and a new factorization lemma. The third contribution is our CGAL-based C++ software for the case of possibly intersecting ellipses, which is the first exact implementation for the problem. Our code spends about a minute to construct the Voronoi diagram of 200 ellipses, when few degeneracies occur. It is faster than the CGAL segment Voronoi diagram, when ellipses are approximated by k-gons for $k > 15$, and a state-of-the-art implementation of the Voronoi diagram of points, when each ellipse is approximated by more than 1250 points.

6.2.5. Patience of Matrix Games

In [15], for matrix games we study how small nonzero probability must be used in optimal strategies. We show that for $n \times n$ win-lose-draw games (i.e. $(-1, 0, 1)$ matrix games) nonzero probabilities smaller than $n^{-O(n)}$ are never needed. We also construct an explicit $n \times n$ win-lose game such that the unique optimal strategy uses a nonzero probability as small as $n^{-\Omega(n)}$. This is done by constructing an explicit $(-1, 1)$ nonsingular $n \times n$ matrix, for which the inverse has only nonnegative entries and where some of the entries are of value $n^{\Omega(n)}$.

6.2.6. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI

Antipodally symmetric spherical functions play a pivotal role in diffusion MRI in representing sub-voxel-resolution microstructural information of the underlying tissue. This information is described by the geometry of the spherical function. In [14] we propose a method to automatically compute all the extrema of a spherical function. We then classify the extrema as maxima, minima and saddle-points to identify the maxima. We take advantage of the fact that a spherical function can be described equivalently in the spherical harmonic (SH) basis, in the symmetric tensor (ST) basis constrained to the sphere, and in the homogeneous polynomial (HP) basis constrained to the sphere. We extract the extrema of the spherical function by computing the stationary points of its constrained HP representation. Instead of using traditional optimization approaches, which are inherently local and require exhaustive search or re-initializations to locate multiple extrema, we use a novel polynomial system solver which analytically brackets all the extrema and refines them numerically, thus missing none and achieving high precision. To illustrate our approach we consider the Orientation Distribution Function (ODF). In diffusion MRI the ODF is a spherical function which represents a state-of-the-art reconstruction algorithm whose maxima are aligned with the dominant fiber bundles. It is, therefore, vital to correctly compute these maxima to detect the fiber bundle directions. To demonstrate the potential of the proposed polynomial approach we compute the extrema of the ODF to extract all its maxima. This polynomial approach is, however, not dependent on the ODF and the framework presented in this line of work can be applied to any spherical function described in either the SH basis, ST basis or the HP basis.

6.2.7. Improving Angular Speed Uniformity by Reparameterization

In [20] we introduce the notion of angular speed uniformity as a quality measure for parameter-izations of plane curves and propose an algorithm to compute uniform reparameterizations for quadratic and cubic curves. We prove that only straight lines have uniform rational parameterizations. For any plane curve other than lines, we show how to find a rational reparameterization that has the maximum uniformity among all the rational parameterizations of the same degree. We also establish specific results for quadratic and certain cubic Bézier curves.

6.2.8. Formalization and Specification of Geometric Knowledge Objects

[7] presents our work on the identification, formalization, structuring, and specification of geometric knowledge objects for the purpose of semantic representation and knowledge management. We classify geometric knowledge according to how it has been accumulated and represented in the geometric literature, formalize geometric knowledge statements by adapting the language of first-order logic, specify knowledge objects with embedded knowledge in a retrievable and extensible data structure, and organize them by modeling the hierarchic structure of relations among them. Some examples of formal specification for geometric knowledge objects are given to illustrate our approach. The underlying idea of the approach has been used successfully for automated geometric reasoning, knowledge base creation, and electronic document generation.

6.2.9. A Framework for Improving Uniformity of Parameterizations of Curves

In [16] we define quasi-speed as a generalization of linear speed and angular speed for parameterizations of curves and use the uniformity of quasi-speed to measure the quality of the parameterizations. With such conceptual setting, a general framework is developed for studying uniformity behaviors under reparameterization via proper parameter transformation and for computing reparameterizations with improved uniformity

of quasispeed by means of optimal single-piece, C^0 piecewise, and C^1 piecewise Möbius transformations. Algorithms are described for uniformity-improved reparameterization using different Möbius transformations with different optimization techniques. Examples are presented to illustrate the concepts, the framework, and the algorithms. Experimental results are provided to validate the framework and to show the efficiency of the algorithms.

6.3. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory

6.3.1. On the Complexity of Solving Quadratic Boolean Systems

A fundamental problem in computer science is to find all the common zeroes of m quadratic polynomials in n unknowns over $\mathbb{F}_2$. The cryptanalysis of several modern ciphers reduces to this problem. Up to now, the best complexity bound was reached by an exhaustive search in $4\log_2 n\, 2^n$ operations. We give an algorithm that reduces the problem to a combination of exhaustive search and sparse linear algebra. This algorithm has several variants depending on the method used for the linear algebra step. Under precise algebraic assumptions on the input system, we show in [4], that the deterministic variant of our algorithm has complexity bounded by $O(2^{0.841n})$ when $m = n$, while a probabilistic variant of the Las Vegas type has expected complexity $O(2^{0.792n})$. Experiments on random systems show that the algebraic assumptions are satisfied with probability very close to 1. We also give a rough estimate for the actual threshold between our method and exhaustive search, which is as low as 200, and thus very relevant for cryptographic applications.

6.3.2. Decomposing polynomial sets into simple sets over finite fields: The positive-dimensional case

Our work in [19] presents an algorithm for decomposing any positive-dimensional polynomial set into simple sets over an arbitrary finite field. The algorithm is based on some relationship established between simple sets and radical ideals, reducing the decomposition problem to the problem of computing the radicals of certain ideals. In addition to direct application of the algorithms of Matsumoto and Kemper, the algorithm of Fortuna and others is optimized and improved for the computation of radicals of special ideals. Preliminary experiments with an implementation of the algorithm in Maple and Singular are carried out to show the effectiveness and efficiency of the algorithm.

6.3.3. Using Symmetries in the Index Calculus for Elliptic Curves Discrete Logarithm

In 2004, an algorithm is introduced to solve the DLP for elliptic curves defined over a non prime finite field $\mathbb{F}_{q^n}$. One of the main steps of this algorithm requires decomposing points of the curve $E(\mathbb{F}_{q^n})$ with respect to a factor base, this problem is denoted PDP. In [11], we apply this algorithm to the case of Edwards curves, the well-known family of elliptic curves that allow faster arithmetic as shown by Bernstein and Lange. More precisely, we show how to take advantage of some symmetries of twisted Edwards and twisted Jacobi intersections curves to gain an exponential factor $2^{\omega(n-1)}$ to solve the corresponding PDP where ω is the exponent in the complexity of multiplying two dense matrices. Practical experiments supporting the theoretical result are also given. For instance, the complexity of solving the ECDLP for twisted Edwards curves defined over $\mathbb{F}_{q^5}$, with $q \approx 2^{64}$, is supposed to be $\sim 2^{160}$ operations in $E(\mathbb{F}_{q^5})$ using generic algorithms compared to 2^{130} operations (multiplication of two 32-bits words) with our method. For these parameters the PDP is intractable with the original algorithm. The main tool to achieve these results relies on the use of the symmetries and the quasi-homogeneous structure induced by these symmetries during the polynomial system solving step. Also, we use a recent work on a new algorithm for the change of ordering of Gröbner basis which provides a better heuristic complexity of the total solving process.

6.3.4. A Distinguisher for High Rate McEliece Cryptosystems

The Goppa Code Distinguishing (GD) problem consists in distinguishing the matrix of a Goppa code from a random matrix. The hardness of this problem is an assumption to prove the security of code-based cryptographic primitives such as McEliece's cryptosystem. Up to now, it is widely believed that the GD

problem is a hard decision problem. We present in [12] the first method allowing to distinguish alternant and Goppa codes over any field. Our technique can solve the GD problem in polynomial-time provided that the codes have sufficiently large rates. The key ingredient is an algebraic characterization of the key-recovery problem. The idea is to consider the rank of a linear system which is obtained by linearizing a particular polynomial system describing a key-recovery attack. Experimentally it appears that this dimension depends on the type of code. Explicit formulas derived from extensive experimentations for the rank are provided for "generic" random, alternant, and Goppa codes over any alphabet. Finally, we give theoretical explanations of these formulas in the case of random codes, alternant codes over any field of characteristic two and binary Goppa codes.

6.3.5. Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic

We investigate in this paper the security of HFE and Multi-HFE schemes as well as their minus and embedding variants. Multi-HFE is a generalization of the well-known HFE schemes. The idea is to use a multivariate quadratic system instead of a univariate polynomial in HFE over an extension field as a private key. According to the authors, this should make the classical direct algebraic (message-recovery) attack proposed by Faugère and Joux on HFE no longer efficient against Multi-HFE. We consider here the hardness of the key-recovery in Multi-HFE and its variants, but also in HFE (both for odd and even characteristic). We first improve and generalize the basic key recovery proposed by Kipnis and Shamir on HFE. To do so, we express this attack as matrix/vector operations. In one hand, this permits to improve the basic Kipnis-Shamir (KS) attack on HFE. On the other hand, this allows to generalize the attack on Multi-HFE. Due to its structure, we prove that a Multi-HFE scheme has much more equivalent keys than a basic HFE. This induces a structural weakness which can be exploited to adapt the KS attack against classical modifiers of multivariate schemes such as minus and embedding. Along the way, we discovered that the KS attack as initially described cannot be applied against HFE in characteristic 2. We have then strongly revised KS in characteristic 2 to make it work. In all cases, the cost of our attacks is related to the complexity of solving MinRank. Thanks to recent complexity results on this problem, we prove that our attack is polynomial in the degree of the extension field for all possible practical settings used in HFE and Multi-HFE. This makes then Multi-HFE less secure than basic HFE for equally-sized keys. As a proof of concept, we have been able to practically break the most conservative proposed parameters of multi-HFE in few days (256 bits security broken in 9 days).

6.3.6. Cryptanalysis of a Public-Key Encryption Scheme Based on New Multivariate Quadratic Assumptions

In [24], we investigate the security of a public-key encryption scheme introduced by Huang, Liu and Yang (HLY) at PKC'12. This new scheme can be provably reduced to the hardness of solving a set of quadratic equations whose coefficients of highest degree are chosen according to a discrete Gaussian distributions. The other terms being chosen uniformly at random. Such a problem is a variant of the classical problem of solving a system of non-linear equations (PoSSo), which is known to be hard for random systems. The main hypothesis of Huang, Liu and Yang is that their variant is not easier than solving PoSSo for random instances. In this paper, we disprove this hypothesis. To this end, we exploit the fact that the new problem proposed by Huang, Liu and Yang reduces to an easy instance of the Learning With Errors (LWE) problem. The main contribution of this paper is to show that security and efficiency are essentially incompatible for the HLY proposal. That is, one cannot find parameters which yield a secure and a practical scheme. For instance, we estimate that a public-key of at least 1.03 GB is required to achieve 80-bit security against known attacks. As a proof of concept, we present practical attacks against all the parameters proposed Huang, Liu and Yang. We have been able to recover the private-key in roughly one day for the first challenge (i.e. Case 1) proposed by HLY and in roughly three days for the second challenge (i.e. Case 2).

6.3.7. On the Complexity of the BKW Algorithm on LWE

In [3], we present 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.

6.3.8. Combined Attack on CRT-RSA. Why Public Verification Must Not Be Public?

In [25] we introduce a new Combined Attack on a CRT-RSA implementation resistant against Side-Channel Analysis and Fault Injection attacks. Such implementations prevent the attacker from obtaining the signature when a fault has been induced during the computation. Indeed, such a value would allow the attacker to recover the RSA private key by computing the gcd of the public modulus and the faulty signature. The principle of our attack is to inject a fault during the signature computation and to perform a Side-Channel Analysis targeting a sensitive value processed during the Fault Injection countermeasure execution. The resulting information is then used to factorize the public modulus, leading to the disclosure of the whole RSA private key. After presenting a detailed account of our attack, we explain how its complexity can be significantly reduced by using Coppersmith's techniques based on lattice reduction. We also provide simulations that confirm the efficiency of our attack as well as two different countermeasures having a very small impact on the performance of the algorithm. As it performs a Side-Channel Analysis during a Fault Injection countermeasure to retrieve the secret value, this article recalls the need for Fault Injection and Side-Channel Analysis countermeasures as monolithic implementations.

6.3.9. Polynomial root finding over local rings and application to error correcting codes

GURUSWAMI and SUDAN designed a polynomial-time list-decoding algorithm. Their method divides into two steps. First it computes a polynomial Q in $\mathbb{F}_q[x][y]$ such that the possible transmitted messages are roots of Q in $\mathbb{F}_q[x]$. In the second step one needs to determine all such roots of Q. Several techniques have been investigated to solve both steps of the problem.

The Guruswami and Sudan algorithm has been adapted to other families of codes such as algebraic-geometric codes and alternant codes over fields. Extensions over certain types of finite rings have further been studied for Reed-Solomon codes, for alternant codes, and for algebraic-geometric codes. In all these cases, the two main steps of the Guruswami and Sudan algorithm are roughly preserved, but to the best of our knowledge, the second step has never been studied into deep details from the complexity point of view. In [5], we investigate root-finding for polynomials over *Galois rings*, which are often used within these error correcting codes, and that are defined as non-ramified extension of $\mathbb{Z}/p^n\mathbb{Z}$. We study the cost of our algorithms, discuss their practical performances, and apply our results to the Guruswami and Sudan list decoding algorithm over Galois rings.

POMDAPI Project-Team (section vide)

PROSECCO Project-Team

6. New Results

6.1. Verification of Security Protocols with Lists in the Symbolic Model

Participants: Bruno Blanchet, Miriam Paiola.

The symbolic model of protocols, or Dolev-Yao model is an abstract model in which messages are represented by terms. Our protocol verifier PROVERIF relies on this model. This year, we have mainly worked on the verification of protocols with lists in this model.

We designed a novel automatic technique for proving secrecy and authentication properties for security protocols that manipulate lists of unbounded length, for an unbounded number of sessions. This result is achieved by extending the Horn clause approach of the automatic protocol verifier ProVerif. We extended the Horn clauses to be able to represent lists of unbounded length. We adapted the resolution algorithm to handle the new class of Horn clauses, and proved the soundness of this new algorithm. We have implemented our algorithm and successfully tested it on several protocol examples, including XML protocols coming from web services. This work has been published in [22] and our prototype is available at http://prosecco.inria.fr/personal/bblanche/publications/BlanchetPaiolaCCS13.html.

Last year, we published a conference paper that shows that, for a limited class of protocols, if a protocol is proven secure by ProVerif with lists of length one, then it is secure for lists of unbounded length. A journal version [50] of this paper has now been accepted.

6.2. Generation of Implementations Proved Secure in the Computational model

Participants: Bruno Blanchet, David Cadé.

The computational model of protocols considers messages as bitstrings, which is more realistic than the formal model, but also makes the proofs more difficult. Our verifier CRYPTOVERIF is sound in this model. This year, we have continued working on our compiler from CRYPTOVERIF speficications to OCaml. Using CryptoVerif and this compiler, we can prove security properties of specifications of protocols in the computational model and generate runnable implementations from such proved specifications. We have published a journal paper on our implementation of SSH generated using this compiler [13] and a proof that this compiler preserves security [23], and we have submitted a journal version of this proof. David Cadé also defended his PhD thesis on this topic [44].

6.3. Computationally Complete Symbolic Attacker and Key Exchange

Participants: Gergely Bana [correspondant], Koji Hasebe, Mitsuhiro Okada.

Around year 2000, various research groups started looking into the relevance of symbolic verification techniques to computational security. If a symbolic verification technique results computational guarantees, we say that computational soundness holds. One of the major concerns has been that the usual Dolev-Yao symbolic attacker that automated symbolic tools used exclusively (to search for attacks) at that time did not seem to allow satisfactory soundness results, only with serious limitations. One possible promising approach to overcome this problem is to derive security guarantees directly as CryptoVerif or F7 does. As an alternative approach, in 2012, Bana and Comon-Lundh introduced a notion they called computationally complete symbolic attacker. With this technique, elimination of the possibility of a computational attack would also mean that computational attack does not exist without the limitations that the Dolev-Yao technique required. Their symbolic attacker can do everything that is not forbidden by conditions derived from standard computational assumptions on the primitives. In this current work, based on predicates for “key compromise”,

we provided such conditions to handle secure encryption even keys are allowed to be sent. We examined both IND-CCA2 and KDM-CCA2 encryptions, both symmetric and asymmetric situations as well as INT-CTXT encryptions. We verified (by hand) a number of protocols as the symmetric Needham-Schroeder protocol, Otway-Rees protocol, Needham-Schroeder-Lowe protocol. Furthermore, we also made some improvements in the computational semantics, and have established a relationship between the computational semantics of Bana and Comon-Lundh and Fitting's embedding of classical logic into S4. This work was published at CCS'13 [19].

6.4. Formal Models and Concrete Attacks on Web Applications

Participants: Karthikeyan Bhargavan [correspondant], Sergio Maffeis, Chetan Bansal, Antoine Delignat-Lavaud, Michael May.

Modern web applications are built as a combination of mostly static servers that host user data and highly dynamic client-side applications that process and present the data to the user. These client-side applications may be hosted as JavaScript within a browser or within custom applications written, say, for smartphones. Hence, in addition to traditional server-side mechanisms, the security of these applications increasingly depends on the correct use of browser-based security mechanisms, client-side access control, and cryptography. These mechanisms are often new, ad hoc, and deserving of close analysis.

Our approach is to formally model various client- and server-side security mechanisms for web applications and rigorously analyze their real-world deployments. When our formal analyses find attacks, we test them against example web applications, report vulnerabilities to various vendors, design countermeasures, and use automated security protocol analysis tools formally verify that our countermeasure resists a large class of attacks. This year, we published three papers in this area. At ESSoS, we formally modeled the authorization policies of common Android apps, found new attacks, and proposed a verified authorization framework [27]. At POST, we formally modeled various cloud-based encrypted storage applications and found both cryptographic and web attacks on them, resulting in patches to these websites and novel countermeasures [20]. At Usenix Security, we proposed a new, safer language for security-critical web components [25]. Defensive JavaScript is a subset of JavaScript that guarantees isolation from other (potentially untrusted) scripts on the same page. This enables, for the first time, the design of cryptographic and single sign-on components that can be formally guaranteed to preserve its secrets even if the hosting website is subject to a cross-site scripting attack.

6.5. Attacks and Proofs for TLS Implementations

Participants: Alfredo Pironti [correspondant], Karthikeyan Bhargavan, Pierre-Yves Strub, Cedric Fournet, Markulf Kohlweiss, Antoine Delignat-Lavaud.

TLS is possibly the most used secure communications protocol, with a 18-year history of flaws and fixes, ranging from its protocol logic to its cryptographic design, and from the Internet standard to its diverse implementations. We have been engaged in a long-term project on verifying TLS implementations and this project is now coming to fruition, with a number of papers are now in the pipeline. We present the main published results below, other papers have been submitted for review.

We have developed a verified reference implementation of TLS 1.2, called miTLS. Our code fully supports its wire formats, ciphersuites, sessions and connections, re-handshakes and resumptions, alerts and errors, and data fragmen- tation, as prescribed in the RFCs; it interoperates with mainstream web browsers and servers. At the same time, our code is carefully structured to enable its modular, automated verification, from its main API down to computational assumptions on its cryptographic algorithms. Our implementation is written in F# and specified in F7. We present security specifications for its main components, such as authenticated stream encryption for the record layer and key establishment for the handshake. We describe their verification using the F7 refinement typechecker. To this end, we equip each cryptographic primitive and construction of TLS with a new typed interface that captures its security prop- erties, and we gradually replace concrete implementations with ideal functionalities. We finally typecheck the protocol state machine, and thus obtain precise security theorems for TLS, as it is implemented and deployed. We also revisit classic attacks and report a few new ones. This work was published at IEEE S&P 2013 [21].

In parallel with this long-term constructive project, we have been analyzing the use of TLS in existing web applications, and our analyses uncovered a number of attacks, leading to patched in popular browsers like Chrome, Internet Explorer, and Firefox, as well as websites like Google and Akamai.

One of these classes of attacks was published at WOOT'13 [29]. In this paper, we identify logical web application flaws which can be exploited by TLS truncation attacks to desynchronize the user- and server-perspective of an application's state. It follows immediately that servers may make false assumptions about users, hence, the flaw constitutes a security vulnerability. Moreover, in the context of authentication systems, we exploit the vulnerability to launch the following practical attacks: we exploit the Helios electronic voting system to cast votes on behalf of honest voters, take full control of Microsoft Live accounts, and gain temporary access to Google accounts.

RAP Project-Team

4. New Results

4.1. Algorithms: Bandwidth Allocation in Optical Networks

Participants: Christine Fricker, Jelena Pesic, Philippe Robert, James Roberts.

The development of dynamic optical switching is widely recognized as an essential requirement to meet anticipated growth in Internet traffic. Since September 2009, RAP has investigated the traffic management and performance evaluation issues that are particular to this technology. Our activity on optical networking is carried out in collaboration with Orange Labs with whom we have a research contract. We have also established contacts with Alcatel-Lucent Bell Labs and had fruitful exchanges with Iraj Saniee and his team on their proposed time-domain wavelength interleaved networking architecture (TWIN).

Our work on access networks proposed an original dynamic bandwidth allocation (DBA) algorithm and demonstrated its excellent performance. This DBA algorithm was then adapted to a meshed metropolitan network based on TWIN and implementing flow-aware resource sharing. Extensions using a concept called “multipath” were shown to offer an energy efficient solution for wide area networks.

In 2013, we contributed to the Celtic Plus project called SASER/SAVENET. This project was approved by the EU in 2012 and funding has been obtained for our participation from the French authorities. The project kickoff meeting was held in November 2012. Our contribution relates to the use of TWIN to create an extended metropolitan optical network. Our partners in the corresponding work package task are Orange, Telecom Bretagne and the engineering school ENSSAT. Overall responsibility for the work package (where alternative optical network architectures are also evaluated) is with Alcatel-Lucent Bell Labs.

In 2013, Inria edited the M12 milestone document of Task 6.4 "TWIN implementations and preliminary MAC protocol specifications". A paper on applying the network architecture and MAC/DBA protocols proposed by the team to the domain of data center interconnects has been submitted.

RAP has continued to work on a two-year research contract with Orange Labs on further developing the multi-path architecture (20012-2013). The main contribution in 2013 has been to propose the use of tunable receivers in addition to tunable transmitters. This technological evolution is possible with recent developments in coherent transmission and offers greater flexibility and enhanced efficiency. Work is continuing on evaluating this architecture by simulation (using Onmet++) and by analytical modelling.

4.2. Algorithms: Content-Centric Networking

Participants: Christine Fricker, Philippe Robert, James Roberts, Nada Sbihi.

RAP participated in an ANR project named CONNECT which contributed to the definition and evaluation of a new paradigm for the future Internet: an information-centric network (ICN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. The project ended in December 2012 but we have continued to work on information-centric networking in 2013.

RAP is participating in an ANR project named CONNECT which contributes to the definition and evaluation of a new paradigm for the future Internet: a content-centric network (CCN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. CCN has been proposed by Van Jacobson and colleagues at the Palo Alto Research Center (PARC). In CCN, content is divided into packet-size chunks identified by a unique name with a particular hierarchical structure. The name and content can be cryptographically encoded and signed, providing a range of security levels. Packets in CCN carry names rather than addresses and this has a fundamental impact on the way the network works. Security concerns are addressed at the content level, relaxing requirements on hosts and

the network. Users no longer need a universally known address, greatly facilitating management of mobility and intermittent connectivity. Content is supplied under receiver control, limiting scope for denial of service attacks and similar abuse. Since chunks are self-certifying, they can be freely replicated, facilitating caching and bringing significant bandwidth economies. CCN applies to both stored content and to content that is dynamically generated, as in a telephone conversation, for example. RAP is contributing to the design of CCN in two main areas:

- the design and evaluation of traffic controls, recognizing that TCP is no longer applicable and queue management will require new, name-based criteria to ensure fairness and to realize service differentiation;
- the design and evaluation of replication and caching strategies that realize an optimal trade-off of expensive bandwidth for cheap memory.

The team also contributes to the development of efficient forwarding strategies and the elaboration of economic arguments that make CCN a viable replacement for IP. CONNECT partners are Alcatel-Lucent (lead), Orange, Inria/RAP, Inria/PLANETE, Telecom ParisTech, UPMC/LIP6.

A paper describing a proposed flow-aware approach for CCN traffic management and its performance evaluation has been presented at the conference Infocom 2012. We have reviewed the literature on cache performance (dating from early work on computer memory management) and identified a practical and versatile tool for evaluating the hit rate (proportion of requests that are satisfied from the cache) as a function of cache size and the assumed object popularity law. This approximate method was first proposed in 2002 by Che, Tung and Wang for their work on web caching. We applied this approximation to evaluate CCN caching performance taking into account the huge population and diverse popularity characteristics that make other approaches ineffective. The excellent accuracy of this method over a wide range of practically relevant traffic models has been explained mathematically. CONNECT ends in December 2012. We are currently defining a new project proposal that should be submitted to the ANR INFRA call in February 2013.

4.3. Scaling Methods: Fluid Limits in Wireless Networks

Participant: Philippe Robert.

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. A fluid scaling analysis of such a network is presented. We have shown that the interaction of several time scales plays an important role in the evolution of such a system, in particular its coordinates may live on very different time and space scales. As a consequence, the associated stochastic processes turn out to have unusual scaling behaviors which give an interesting fairness property to this class of algorithms. A heavy traffic limit theorem for the invariant distribution has also been proved. A generalization to the resource sharing algorithm for which the $\log$ function is replaced by an increasing function.

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. Stochastic Modeling of Biological Networks

Participants: Emanuele Leoncini, Philippe Robert.

This is a collaboration with Vincent Fromion from INRA Jouy en Josas, which started on October 2010.

The goal is to propose a mathematical model of the production of proteins in prokaryotes. Proteins are biochemical compounds that play a key role in almost all the cell functions and are crucial for cell survival and for life in general. In bacteria the protein production system has to be capable to produce abut 2500 different types of proteins in different proportions (from few dozens for the replication machinery up to 100000 for certain key metabolic enzymes). Bacteria uses more than the 85% of their resources to the protein production, making it the most relevant process in these organisms. Moreover this production system must meet two opposing problems: on one side it must provide a minimal quantity for each protein type in order to ensure the smooth-running of the cell, on the other side an "overproduction policy" for all the proteins is infeasible, since this would impact the global performance of the system and of the bacterium itself.

Gene expression is intrinsically a stochastic process: gene activation/deactivation occurs by means the encounter of polymerase/repressor with the specific gene, moreover many molecules that take part in the protein production act at extremely low concentrations. We have restated mathematically the classical model using Poisson point processes (PPP). This representation, well-known in the field of queueing networks but, as far as we know, new in the gene expression modeling, allowed us to weaken few hypothesis of the existing models, in particular the Poisson hypothesis, which is well-suited in some cases, but that, in some situations, is far from the biological reality as we consider for instance the protein assemblage.

The theoretical environment of Poisson point processes has lead us to propose a new model of gene expression which captures on one side the main mechanisms of the gene expression and on the other side it tries to consider hypothesis that are more significant from a biological viewpoint. In particular we have modeled: gene activation/deactivation, mRNA production and degradation, ribosome attachment on mRNA, protein elongation and degradation. We have shown how the probability distribution of the protein production and the protein lifetime may have a significant impact on the fluctuations of the number of proteins. We have obtained analytic formulas when the duration of protein assemblage and degradation follows a general probability distribution, i.e. without the Poisson hypothesis. In particular, by using a PPP representation we have been able to include the deterministic continuous phenomenon of protein degradation, which is the main protein degradation mechanism for stable proteins. We have showed moreover that this more realistic description is surprisingly identical in distribution with the classic assumption of protein degradation by means of a degrading protein (*proteosome*). We have used our model also to compare the variances resulting by choosing different hypotheses for the probability elongation, in particular we have hypothesize the protein assembly to be deterministic. This assumption is justified because of the elongation step, which consists of a large number of elementary steps, can be described by the sum of exponential steps and the resulting distribution is well approximated by a Gaussian distribution because of the central limit theorem. Under the hypothesis of small variance of the resulting Gaussian distribution, we can assume the elongation step to be deterministic. The model has showed how, under the previous hypothesis, the variance on the number of proteins is bigger than the classical model with the Poisson hypothesis.

We have developed a C++ stochastic simulator for our general model, which has allowed the computation of variance when it was not possible to derive explicit analytic close formulas and the simulation of some extension of the actual model.

This year we have investigated a mathematical model of the production of proteins in prokaryotic cells. Up to now most of the mathematical used to study these problems concern the production of *one* fixed class of proteins. When several classes of proteins are considered, each class requires in fact a fraction of the common and limited resources of the cell. One has therefore to understand how the allocation of the resources within the cell is done. Due to the fact that the cytoplasm of the cell is a quite disorganized medium where the components of the cell move, the whole production process has an important stochastic component. A model describing the allocation of the ribosomes of the cell to produce proteins is investigated via a Markovian representation. Asymptotic results for the equilibrium and for the transient behavior have been obtained under a scaling procedure and a reasonable biological assumption of saturation, i.e. when resources of the cell are tight. The equilibrium and the transient behavior have been investigated, it has been shown in particullar that, in the limit, the number of free ribosomes converges in distribution to a Poisson distribution whose parameter satisfies a fixed point equation.

4.5. Stochastic networks: large bike sharing systems

Participants: Christine Fricker, Hanène Mohamed, Nicolas Servel.

This is a collaboration with Nicolas Gast (EPFL). Bike sharing systems were launched by numerous cities to be a urban mode of transportation, for example Velib in Paris. One of the major issues is the availability of the resources: bikes or free slots to return the bikes. These systems became a hot topic in Operation Research and now the importance of stochasticity of such system behavior is commonly admitted. The problem is to understand their behavior and how to manage them in order to provide both resources to users.

Our model is the first one taking into account the finite number of spots at the stations. In a homogeneous model, mean field limit theorems give the dynamic of a large system. Analytical results are obtained and convergence proved in a standard model via Lyapunov functions. It allows to find the best ratio of bikes per station and to measure the improvement of incentive mechanisms, as choosing among two stations for example. We investigate also redistribution of bikes by trucks. Further results deal with heterogeneous system. By mean field techniques, analytical results were recently obtained on systems consisting in several clusters. In a work with Nicolas Servel, we discuss the improvement of choosing between two stations in the same cluster. Our goal is to propose, via a theoretical study and tests, simple algorithms to improve the system behavior.

With Hanene Mohamed, we study the problem of impact of geometry on incentive mechanisms. Our first model under investigation is very close from the Gates-Westcott crystal growth model with its underlying random deposition process.

4.6. Random Graphs

Participants: Nicolas Broutin, Henning Sulzbach.

4.6.1. Connectivity in models of wireless networks

This is joint work with S. Boucheron (Paris 7), L. Devroye (McGill), N. Fraiman (McGill), and G. Lugosi (Pompeu Fabra).

The traditional models for wireless networks rely on geometric random graphs. However, if one wants to ensure that the graph be fully connected the radius of influence (hence the power necessary, and number of links) is too large to be fully scalable. Recently some models have been proposed that skim the neighbours and only retain a random subset for each node, hence creating a sparser overlay that would hopefully be more scalable. The first results on the size of the subsets which guarantee connectivity of overlay (the irrigation graph) confirm that the average number of links per node is much smaller, but it remains large. These results motivate further investigations on the size of the largest connected component when one enforces a constant average degree which are in the process of being written.

4.6.2. Random graphs and minimum spanning trees

This is a long term collaboration with L. Addario-Berry (McGill), C. Goldschmidt (Oxford) and G. Miermont (ENS Lyon).

The random graph of Erdős and Rényi is one of the most studied models of random networks. Among the different ranges of density of edges, the "critical window" is the most interesting, both for its applications to the physics of phase transitions and its applications to combinatorial optimization (minimum spanning tree, constraint satisfaction problems). One of the major questions consists in determining the distribution of distances between the nodes. A limit object (a scaling limit) has been identified, that allows to describe precisely the first order asymptotics of pairwise distances between the nodes. This limit object is a random metric space whose definition allows to exhibit a strong connection between random graphs and the continuum random tree of Aldous. A variety of questions like the diameter, the size of cycles, etc, may be answered immediately by reading them on the limit metric space.

In a stochastic context, the minimum spanning tree is tigthly connected to random graphs via Kruskal's algorithm. Random minimum spanning trees have attracted much research because of their importance in combinatorial optimization and statistical physics; however, until now, only parameters that can be grasped by local arguments had been studied. The scaling limit of the random graphs obtained permits to describe precisely the metric space scaling limit of a random minimum spanning tree, which identifies a novel continuum random tree which is truely different from that of Aldous.

4.6.3. Analysis of recursive partitions

This is joint work with R. Neininger (Frankfurt)

The techniques that we developped in order to estimate the cost of partial match queries in random quad trees have been used to solve an open question about the recursive lamination of the disk. We have proved that the planar dual of the lamination, which is a tree, converges almost surely when suitably rescaled to a compact random tree encoded by a continuous function. We also pinned down the fractal dimension of the limit object.

REGAL Project-Team

5. New Results

5.1. Introduction

In 2013, we focused our research on the following areas:

- *Distributed algorithms for dynamic and large networks.*
- *Management of distributed data.*
- *Performance and robustness of Systems Software in multicore architectures.*

5.2. Distributed algorithms for dynamic networks

Participants: Luciana Bezerra Arantes [correspondent], Rudyar Cortes, Guthemberg Da Silva Silvestre, Raluca Diaconu, Ruijing Hu, Anissa Lamani, Jonathan Lejeune, Olivier Marin, Sébastien Monnet, Franck Petit [correspondent], Karine Pires, Maria Potop-Butucaru, Pierre Sens, Véronique Simon, Julien Sopena.

This objective aims to design distributed algorithms adapted to new large scale or dynamic distributed systems, such as mobile networks, sensor networks, P2P systems, Grids, Cloud environments, and robot networks. Efficiency in such demanding environments requires specialised protocols, providing features such as fault or heterogeneity tolerance, scalability, quality of service, and self-stabilization. Our approach covers the whole spectrum from theory to experimentation. We design algorithms, prove them correct, implement them, and evaluate them in simulation, using OMNeT++ or PeerSim, and on large-scale real platforms such as Grid'5000. The theory ensures that our solutions are correct and whenever possible optimal; experimental evidence is necessary to show that they are relevant and practical.

Within this thread, we have considered a number of specific applications, including massively multi-player on-line games (MMOGs) and peer certification.

Since 2008, 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: mutual exclusion, failure detection, data dissemination and data finding in large scale systems, self-stabilization and self-* services.

5.2.1. Mutual Exclusion and Failure Detection.

Mutual Exclusion and Fault Tolerance are two major basic building blocks in the design of distributed systems. Most of the current mutual exclusion algorithms are not suitable for modern distributed architectures because they are not scalable, they ignore the network topology, and they do not consider application quality of service constraints. Under the ANR Project *MyCloud* and the FSE *Nu@age*, we study locking algorithms fulfilling some QoS constraints often found in Cloud Computing [46], [38].

A classical way for a distributed system to tolerate failures is to detect them and then recover. It is now well recognized that the dominant factor in system unavailability lies in the failure detection phase. Regal has worked for many years on practical and theoretical aspects of failure detections and pioneered hierarchical scalable failure detectors. [2] Since 2008, we have studied the adaptation of failure detectors to dynamic networks. In 2013, we studied Ω, the eventual leader election failure detector. Ω ensures that, eventually, each process in the system will be provided by an unique leader, elected among the set of correct processes in spite of crashes and uncertainties. It is known to be weakest failure detector to solve agreement protocols such as Paxos. Then, a number of eventual leader election protocols were suggested. Nonetheless, as far as we are aware of, no one of these protocols tolerates a free pattern of node mobility. In [27] we propose a new protocol for this scenario of dynamic and mobile unknown networks.

[2] Recent work by Leners et al published in SOSP 2011 uses our DSN 2003 paper as basis for performance comparison

5.2.2. Self-Stabilization and Self-* Services.

We have also approached fault tolerance through self-stabilization. Self-stabilization is a versatile technique to design distributed algorithms that withstand transient faults. In particular, we have worked on the unison problem, [3] i.e., the design of self-stabilizing algorithms to synchronize a distributed clock. As part of the ANR project *SPADES*, we have proposed several snap-stabilizing algorithms for the message forwarding problem that are optimal in terms of number of required buffers. A snap-stabilizing algorithm is a self-stabilizing algorithm that stabilizes in 0 steps; in other words, such an algorithm always behaves according to its specification.

Finally, we have applied our expertise in distributed algorithms for dynamic and self-* systems in domains that at first glance seem quite far from the core expertise of the team, namely ad-hoc systems and swarms of mobile robots. In the latter, as part of ANR project *R-Discover*, we have studied various problems such as exploration and gathering.

5.2.3. Dissemination and Data Finding in Large Scale Systems.

In the area of large-scale P2P networks, we have studied the problems of data dissemination and overlay maintenance, i.e., maintenance of a logical network built over the a P2P network. In 2013, we have proposed a new distributed algorithm suitable for scale-free random topologies which model some complex real world networks [37], [52].

5.2.4. Peer certification.

In a distributed system, the certification of transactions makes it possible to circumscribe malicious behaviors. Certification requires the use of a trusted third party which must be centralized to guarantee safety. At a large scale, however, centralized certification represents a bottleneck and a single point of attack or failure.

We proposed two decentralized approaches towards certifying transactions with a high probability of success. The first approach replicates transactions over multiple peers and retains identical results from a qualified majority to certify that a service has been carried out for a given client at a given time [30]. The second approach uses distributed reputations to identify trusted nodes and use them as game referees to detect and prevent cheating [57].

5.3. Management of distributed data

Participants: Pierpaolo Cincilla, Guthemberg Da Silva Silvestre, Raluca Diaconu, Jonathan Lejeune, Mesaac Makpangou, Olivier Marin, Sébastien Monnet, Dang Nhan Nguyen, Burcu Külahçioglu Özkan, Karine Pires, Masoud Saeida Ardekani, Thomas Preud'Homme, Pierre Sens, Marc Shapiro, Véronique Simon, Julien Sopena, Gaël Thomas, Mathieu Valero, Mudit Verma, Marek Zawirski.

Storing and sharing information is one of the major reasons for the use of large-scale distributed computer systems. Replicating data at multiple locations ensures that the information persists despite the occurrence of faults, and improves application performance by bringing data close to its point of use, enabling parallel reads, and balancing load. This raises numerous issues:

- where to store or replicate the data, in order to ensure that it is available quickly and remains persistent despite failures and disconnections;
- how many copies are needed to face dynamically-changing demand (load) and offer (elasticity);
- how to parallelize writes and hence how to ensure consistency between replicas;
- tradeoffs between synchronised, consistent but slow updates, and fast but weakly-consistent ones;
- when and how to move data to computation, or computation to data, in order to improve response time while minimizing storage or energy usage;
- etc.

[3] C. Boulinier, F. Petit, and V. Villain. Synchronous vs. asynchronous unison. Algorithmica, 51(1):61-80, 2008

5.3.1. Long term durability

To tolerate failures, distributed storage systems replicate data. However, despite the replication, pieces of data may be lost (i.e. all the copies are lost). We have previously proposed a mechanism, RelaxDHT, to make distributed hash tables (DHT) resilient to high churn rates.

Well sized systems rarely loose data, still, data may be lost: the more the time passes, the greater is the risk of loss. It is thus necessary to study data durability on a long term. To do so, we have implemented an efficient simulator, we can simulate a 100 node system over years within several hours. We have observe that a given system with a given replication mechanism can store a certain amount of data above which the loss rate would be greater than an "acceptable"/fixed threshold. This amount of data can be used as a metric to compare replication strategies. We have studied the impact of the data distribution layout upon the loss rate. The way the replication mechanism distribute the data copies among the nodes has a great impact. If node contents are very correlated, the number of available sources to heal a failure is low. On the opposite, if the data copies are shuffled among the nodes, many source nodes may be available to heal the system, and thus, the system losses less pieces of data. We are also studying the impact of other parameters, like the replication degree or the way a new storer node is chosen.

5.3.2. Adaptative replication

Different pieces of data have different popularity: some data are stored but never accessed while other pieces are very "hot" and are requested concurrently by many clients. This implies that different pieces of data with different popularity should have a different number of copies to efficiently serve the requests without wasting resources. Furthermore, for a given piece of data, the popularity may vary drastically among time. It is thus important that the replication mechanism dynamically adapt the number of replicas to the demand. In the context of the ODISEA2 FUI project, we have made two main contributions. First, we have studied the popularity distribution and evolution of live video streams (Karine Pires thesis). Second, we have designed replication mechanisms able to gracefully adapt the replication degree to the demand, one based on bandwidth reservation, and one using statistical learning (Guthemberg Silvestre thesis).

5.3.3. Strong consistency

When data is updated somewhere on the network, it may become inconsistent with data elsewhere, especially in the presence of concurrent updates, network failures, and hardware or software crashes. A primitive such as consensus (or equivalently, total-order broadcast) synchronises all the network nodes, ensuring that they all observe the same updates in the same order, thus ensuring strong consistency. However the latency of consensus is very large in wide-area networks, directly impacting the response time of every update. Our contributions consist mainly of leveraging application-specific knowledge to decrease the amount of synchronisation.

To reduce the latency of consensus, we study *Generalised Consensus* algorithms, i.e., ones that leverage the commutativity of operations or the spontaneous ordering of messages by the network. We propose a novel protocol for generalised consensus that is optimal, both in message complexity and in faults tolerated, and that switches optimally between its fast path (which avoids ordering commuting requests) and its classical path (which generates a total order). Experimental evaluation shows that our algorithm is much more efficient and scales better than competing protocols.

When a database is very large, it pays off to replicate only a subset at any given node; this is known as partial replication. This allows non-overlapping transactions to proceed in parallel at different locations and decreases the overall network traffic. However, this makes it much harder to maintain consistency. We designed and implemented two *genuine* consensus protocols for partial replication, i.e., ones in which only relevant replicas participate in the commit of a transaction.

Another research direction leverages isolation levels, particularly Snapshot Isolation (SI), in order to parallelize non-conflicting transactions on databases. We prove a novel impossibility result: under standard assumptions (data store accesses are not known in advance, and transactions may access arbitrary objects in the

data store), it is impossible to have both SI and GPR. Our impossibility result is based on a novel decomposition of SI which proves that, like serializability, SI is expressible on plain histories. These results are published at the Euro-Par conference [42].

We designed an efficient protocol that maintains side-steps this impossibility but maintains the most important features of SI:

1. (Genuine Partial Replication) only replicas updated by a transaction T make steps to execute T;
2. (Wait-Free Queries) a read-only transaction never waits for concurrent transactions and always commits;
3. (Minimal Commit Synchronization) two transactions synchronize with each other only if their writes conflict.

The protocol also ensures Forward Freshness, i.e., that a transaction may read object versions committed after it started.

Non-Monotonic Snapshot Isolation (NMSI) is the first strong consistency criterion to allow implementations with all four properties. We also present a practical implementation of NMSI called Jessy, which we compare experimentally against a number of well-known criteria. Our measurements show that the latency and throughput of NMSI are comparable to the weakest criterion, read-committed, and between two to fourteen times faster than well-known strong consistencies. This was published in the Symp. on Reliable Distr. Sys. (SRDS) [43].

5.3.4. Distributed Transaction Scheduling

Parallel transactions in distributed DBs incur high overhead for concurrency control and aborts. Our Gargamel system proposes an alternative approach by pre-serializing possibly conflicting transactions, and parallelizing non-conflicting update transactions to different replicas. This system provides strong transactional guarantees. In effect, Gargamel partitions the database dynamically according to the update workload. Each database replica runs sequentially, at full bandwidth; mutual synchronisation between replicas remains minimal. Our simulations show that Gargamel improves both response time and load by an order of magnitude when contention is high (highly loaded system with bounded resources), and that otherwise slow-down is negligible.

Our current experiments aim to compare the practical pros and cons of different approaches to designing large-scale replicated databases, by implementing and benchmarking a number of different protocols.

5.3.5. Eventual consistency

Eventual Consistency (EC) aims to minimize synchronisation, by weakening the consistency model. The idea is to allow updates at different nodes to proceed without any synchronisation, and to propagate the updates asynchronously, in the hope that replicas converge once all nodes have received all updates. EC was invented for mobile/disconnected computing, where communication is impossible (or prohibitively costly). EC also appears very appealing in large-scale computing environments such as P2P and cloud computing. However, its apparent simplicity is deceptive; in particular, the general EC model exposes tentative values, conflict resolution, and rollback to applications and users. Our research aims to better understand EC and to make it more accessible to developers.

We propose a new model, called *Strong Eventual Consistency* (SEC), which adds the guarantee that every update is durable and the application never observes a roll-back. SEC is ensured if all concurrent updates have a deterministic outcome. As a realization of SEC, we have also proposed the concept of a Conflict-free Replicated Data Type (CRDT). CRDTs represent a sweet spot in consistency design: they support concurrent updates, they ensure availability and fault tolerance, and they are scalable; yet they provide simple and understandable consistency guarantees.

This new model is suited to large-scale systems, such as P2P or cloud computing. For instance, we propose a “sequence” CRDT type called Treedoc that supports concurrent text editing at a large scale, e.g., for a wikipedia-style concurrent editing application. We designed a number of CRDTs such as counters (supporting concurrent increments and decrements), sets (adding and removing elements), graphs (adding and removing vertices and edges), and maps (adding, removing, and setting key-value pairs).

On the theoretical side, we identified sufficient correctness conditions for CRDTs, viz., that concurrent updates commute, or that the state is a monotonic semi-lattice. CRDTs raise challenging research issues: What is the power of CRDTs? Are the sufficient conditions necessary? How to engineer interesting data types to be CRDTs? How to garbage collect obsolete state without synchronisation, and without violating the monotonic semi-lattice requirement? What are the upper and lower bounds of CRDTs? We co-authored an innovative approach to these questions, to be published at Principles of Programming Languages (POPL) 2014 [29].

We are currently developing an extreme-scale CRDT platform called SwiftCloud; see Section 4.2 .

5.3.6. Mixing commutative and non-commutative updates: reservations

Asynchronous updates are desirable because they ensure the system is available, fast and scalable. CRDTs are asynchronous, but cannot guarantee strong invariants, such as ensuring that a shared counter never goes negative. To solve this problem, we define a novel hybrid model that supports both synchronous and asynchronous updates, "red-blue-purple" consistency. The RPB model classifies updates into commutative, partially-commutative and non-commutative, and distinguishes the (global) states where partially-commutative operations can safely run asynchronously. We use reservation techniques to ensure operation in such states. A reservation promises, to a cache that holds it, that the system is in a state that allows the cache server to perform purple updates asynchronously. Reservations ensure that data is in a known state by caching both data and access permissions over data to make updates. This approach strengthens the safety guarantees in addition to eventual consistency [40].

5.4. Performance and Robustness of Systems Software in Multicore Architectures

Participants: Koutheir Attouchi, Harris Bakiras, Antoine Blin, Florian David, Bertil Folliot, Lokesh Gidra, Julia Lawall, Jean-Pierre Lozi, Gilles Muller [correspondent], Dang Nhan Nguyen, Thomas Preud'Homme, Suman Saha, Peter Senna Tschudin, Marc Shapiro, Julien Sopena, Gaël Thomas, Mudit Verma.

5.4.1. Managed Runtime Environments

Today, multicore architectures are becoming ubiquitous, found even in embedded systems, and thus it is essential that managed runtime environments can scale on multicore processors. We have found that two major scalability bottlenecks are the implementation of highly contented locks and of garbage collectors. On a multicore, a single lock can overload the bus because the cache line that contains the lock bounces between the cores, eliminating all the performance benefits from adding more cores. To address this issue, as part of the PhD of Jean-Pierre Lozi, we have developed remote core locking (RCL), in which highly contended locks are implemented on a dedicated server, minimizing bus traffic and improving application scalability. This work initially targeted C code but is now being adapted to the needs of Java applications in the PhD of Florian David. For garbage collectors, as the memory is physically distributed among a set of memory controllers, a collection saturates the bus when the collector threads access remote memory. This saturation prevents the garbage collector from scaling with the number of cores, making the garbage collector a major bottleneck of managed runtime environments on multicore hardware. As part of the PhD of Lokesh Gidra, we have identified memory placement schemes that decrease the number of remote memory accesses during a collection in OpenJDK 7, thus preventing the bottleneck caused by bus saturation [36].

5.4.2. System software robustness

A widely recognized problem in the area of finding bugs in API usage in systems code is to know what APIs are expected and to identify contexts where these expectations are not satisfied. Indeed, systems code, such as an operating systems kernel, is typically voluminous, amounting to millions of lines of code, and uses many different highly specialized APIs, making it impossible for most developers to keep the usage protocols of all of them in mind. To address this issue, we have developed an approach to inferring API function usage protocols from software, relying on knowledge of common code structures (Software – Practice and Experience [26]). Building on this experience, we have developed an approach to finding resource-release omission faults in

systems code that leverages information local to a single function [44]. This approach permits finding hundreds of faults in Linux kernel code as well as a variety of other systems software, with a low rate of false positives. Finally, we have initiated an effort on understanding the range and scope of the oops reports collected in the recently revived Linux kernel oops repository [59].

Beyond finding faults in existing code, we have also considered how systems code is constructed. Specifically, in the context of Linux device drivers, we have identified the notion of a *gene*, as a sequence of code fragments that express a particular device or operating system functionality. We have performed an initial partial sequencing of the genes making up the probe functions of Linux platform drivers [45]. Relatedly, in the context of a Merlion collaboration grant with David Lo of Singapore Management University, we have considered the problem of recommending APIs to developers. We propose one approach based on the set of libraries used by other software having similar properties [47], and a second approach based on the set of libraries used to implement related feature requests [48].

5.4.3. Domain-specific languages for systems software

A challenge in the management of a datacenter is the placement of application replicas, both to avoid a single point of failure and to limit communication costs. We have proposed a novel approach, BtrPlace [23], based on the use of a domain-specific language to express constraints derived from properties of the application and of the datacenter, and the use of a constraint solver to efficiently resolve these constraints. Simulations show that BtrPlace is able to repair a configuration involving 5000 servers after a server failure in 3 minutes.

While the use of domain-specific languages such as that of BtrPlace can ease programming, it is well known that developing, and especially maintaining, a domain-specific language over time is time-consuming and challenging. This is particularly the case when the domain-specific language provides domain-specific verifications, as the code implementing these verifications has to be maintained along with the rest of the language implementation. Furthermore, new domain-specific languages typically must evolve frequently, as the language developer comes to better understand the range and scope of the domain. To address these issues, we have proposed a methodology for domain-specific language implementation development for C-like domain-specific languages [19], based on the use of rewriting rules implemented using Coccinelle. We apply this approach to our previously developed domain specific language z2z for developing network gateways, and find that the resulting language implementation is more concise and easier to extend with new language features.

REO Project-Team

6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Mikel Landajuela Larma, Jimmy Mullaert, Marina Vidrascu.

- In [17] we analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – to estimate the state of a fluid-structure interaction model for hemodynamics, when the measurements are assumed to be restricted to displacements or velocities in the solid. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect per se. Whereas improvements can be sought for when additional measurements are available in the fluid domain, the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters.
- In [28] we introduce a class of incremental displacement-correction schemes for the explicit coupling of a thin-structure with an incompressible fluid. We provide a general stability and convergence analysis that covers both the incremental and the non-incremental variants. Their stability properties are independent of the added-mass effect. The superior accuracy of the incremental schemes (with respect to the original non-incremental variant) is highlighted by the error estimates, and then confirmed in a benchmark by numerical experiments.
- In [29], [62] we introduce a class of fully decoupled time-marching schemes (velocity-pressure-displacement splitting) for the coupling of an incompressible fluid with a thin-walled viscoelastic structure. A priori energy estimates guaranteeing unconditional stability are established for the variants without extrapolation and with first-order extrapolation. The accuracy and performance of the methods proposed are discussed in several numerical examples.
- In [30] we introduce a class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and a general thin-walled structure (e.g., including damping and non-linear behavior). The fundamental ingredient in these methods is the (parameter free) explicit Robin interface condition for the fluid, which enables the fluid-solid splitting through appropriate extrapolations of the solid velocity and fluid stress on the interface. The resulting solution procedures are genuinely partitioned. Stability and error estimates are provided for all the variants (depending on the extrapolations), using energy arguments within a representative linear setting. In particular, we show that one of them yields added-mass free unconditional stability and optimal (firs-order) time accuracy. A comprehensive numerical study, involving different examples from the literature, supports the theory.
- In [63] we introduce a new class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and an elastic structure. These methods generalize the arguments reported in [28], [30] to the case of the coupling with thick-walled structures. The basic idea lies in the derivation of an intrinsic interface Robin consistency at

the space semi-discrete level, using a lumped-mass approximation in the structure. The fluid-solid splitting is then performed through appropriate extrapolations of the solid velocity and stress on the interface. Based on these methods, a new, parameter-free, Robin-Neumann iterative procedure is also proposed for the partitioned solution of implicit coupling. A priori energy estimates, guaranteeing the (added-mass free) stability of the schemes and the convergence of the iterative procedure, are established. The accuracy and robustness of the methods are illustrated in several numerical examples.

- In [22] we discuss explicit coupling schemes for fluid-structure interaction problems where the added mass effect is important. We show the close relation between coupling schemes using Nitsche's method and a Robin-Robin type coupling. In the latter case the method may be implemented either using boundary integrals of the stresses or the more conventional discrete lifting operators. We also make the observation that these scheme are stable under a hyperbolic type CFL condition, but that optimal accuracy imposes a parabolic type CFL conditions due to the splitting error. Two strategies to enhance the accuracy of the coupling scheme under the hyperbolic CFL-condition are suggested, one using extrapolation and defect-correction and one using a penalty-free non-symmetric Nitsche method. Finally we illustrate the performance of the proposed schemes on some numerical examples in two and three space dimensions.
- In [60] we consider the extension of the Nitsche method to the case of fluid-structure interaction problems on unfitted meshes. We give a stability analysis for the space semi-discretized problem and show how this estimate may be used to derive optimal error estimates for smooth solutions, irrespective of the mesh/interface intersection. Some numerical examples illustrate the theoretical discussion.
- In [21] we are interested by the three-dimensional coupling between an incompressible fluid and a rigid body. The fluid is modeled by the Navier-Stokes equations, while the solid satisfies the Newton's laws. In the main result of the paper we prove that, with the help of a distributed control, we can drive the fluid and structure velocities to zero and the solid to a reference position provided that the initial velocities are small enough and the initial position of the structure is close to the reference position. This is done without any condition on the geometry of the rigid body.
- In the book chapter [57] we deal with some specific existence and numerical results applied to a 2D/1D fluid-structure coupled model, for an incompressible fluid and a thin elastic structure. We underline some of the mathematical and numerical difficulties that one may face when studying this kind of problems such as the geometrical nonlinearities or the added mass effect. In particular we underline the link between the strategies of proof of weak or strong solutions and the possible algorithms to discretize these types of coupled problems.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Grégory Arbia, Benoit Fabrèges, Jean-Frédéric Gerbeau, Sanjay Pant, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [61] we propose a new approach to the loosely coupled time-marching of a fluid-fluid interaction problems involving the incompressible Navier-Stokes equations. The methods combine a specific explicit Robin-Robin treatment of the interface coupling with a weakly consistent interface pressure stabilization in time. A priori energy estimates guaranteeing stability of the splitting are obtained for a total pressure formulation of the coupled problem. The performance of the proposed schemes is illustrated on several numerical experiments related to simulation of aortic blood flow.
- In [49] we present our strategy to meet the MICCAI Challenge 2013: the goal was to recover a measured (but unrevealed) pressure drop across a coarctation of the aorta through 3D simulation. A filtering-based strategy is devised to perform parameter estimation and subsequent multiscale CFD simulations of arterial blood flow. The method is applied to the patient-specific case in the two physiological states of rest and stress. In both cases, the method is shown to be effective in closely matching the available clinically measured data. Pressure drop across the coarctation is predicted

for both states. At the time of [47], these measurements were available: the computed pressure drop across the coarctation for the stress case appears to be very close to the measured one, while the one for the rest case is not as good. One should note that no participant of the challenge managed to recover the measured pressure drop for the rest case.

- In [35], we aim to reduce the complexity of patient-specific simulations by combining image analysis, computational fluid dynamics and model order reduction techniques. The proposed method makes use of a reference geometry estimated as an average of the population, within an efficient statistical framework based on the currents representation of shapes. Snapshots of blood flow simulations performed in the reference geometry are used to build a POD (Proper Orthogonal Decomposition) basis, which can then be mapped on new patients to perform reduced order blood flow simulations with patient specific boundary conditions. This approach is applied to a data-set of 17 tetralogy of Fallot patients to simulate blood flow through the pulmonary artery under normal (healthy or synthetic valves with almost no backflow) and pathological (leaky or absent valve with backflow) conditions to better understand the impact of regurgitated blood on pressure and velocity at the outflow tracts. The model reduction approach is further tested by performing patient simulations under exercise and varying degrees of pathophysiological conditions based on reduction of reference solutions (rest and medium backflow conditions respectively).
- In [16], we analyze two 3D-0D coupling approaches in which a fractional-step projection scheme is used in the fluid. Our analysis shows that explicit approaches might yield numerical instabilities, particularly in the case of realistic geometries with multiple outlets. We introduce and analyze an implicitly 3D-0D coupled formulation with enhanced stability properties and which requires a negligible additional computational cost. Furthermore, we also address the extension of these methods to fluid-structure interaction problems. The theoretical stability results are confirmed by meaningful numerical experiments in patient specific geometries coming from medical imaging.
- In [36], we developed two multi-scale models, each including the 3D model of the surgical junction constructed from MRI, and a closed-loop LPN derived from pre-operative data obtained from two patients prior to Stage 2 Fontan palliation of single ventricle congenital heart disease. "Virtual" surgeries were performed and a corresponding multi-scale simulation predicted the patient's post-operative hemodynamic conditions, tested under different physiological conditions. The impact of the surgical junction geometry on the global circulation was contrasted with variations of key physiological parameters.
- A novel Y-shaped baffle was proposed for the Stage 3 Fontan operation achieving overall superior hemodynamic performance compared with traditional designs. Previously, we investigated if and how the inferior vena cava flow (which contains an important biological hepatic factor) could be best distributed among both lungs. In [43] we proposed a multi-step method for patient-specific optimization of such surgeries to study the effects of boundary conditions and geometry on hepatic factor distribution (HFD). The resulting optimal Y-graft geometry largely depended on the patient left/right pulmonary flow split. Unequal branch size and constrained optimization on energy efficiency were explored. Two patient-specific examples showed that optimization-derived Y-grafts effectively improved HFD.
- The use of elaborate closed-loop lumped parameter network (LPN) models of the heart and the circulatory system as boundary conditions for 3D simulations can provide valuable global dynamic information, particularly for patient specific simulations. In [27], we have developed and tested a numerical method to couple a 3D Navier-Stokes finite-element formulation and a reduced model of the rest of the circulation, keeping the coupling robust but modular. For Neumann boundaries, implicit, semi-implicit, and explicit quasi-Newton formulations are compared within the time-implicit coupling scheme. The requirements for coupling Dirichlet boundary conditions are also discussed and compared to that of the Neumann coupled boundaries. Both these works were key for applications where blood flows in different directions during the cardiac cycle and where coupling with the rest of the circulation is instrumental (see the shunt optimization application [75]).

- In [26] we propose the first patient-specific predictive modeling of stage 2 palliation for congenital heart disease by using virtual surgery and closed-loop multi-scale modeling. We present a workflow to perform post-operative simulations from pre-operative clinical data. Two surgical options (bi-directional Glenn and hemi-Fontan operations) are virtually performed and coupled to the pre-operative LPM, with the hemodynamics of both options reported. Results are validated against postoperative clinical data.
- In [14] we study the influence of solvers and test case setup on the result of numerical simulations. The need for detailed construction of the numerical model depends on the precision needed to answer the biomedical question at hand and should be assessed for each problem on a combination of clinically relevant patient-specific geometry and physiological conditions. Methods and results between a commercial code and an in-house research code are illustrated on three congenital heart disease examples of increasing complexity. This publication is designed as a tool to better communicate with clinical researchers interested in simulations.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Annabelle Collin, Elisa Schenone.

- In [25], Computational electrophysiology is a very active field with tremendous potential in medical applications, albeit leads to highly intensive simulations. We here propose a surface-based electrophysiology formulation, motivated by the modeling of thin structures such as cardiac atria, which greatly reduces the size of the computational models. Moreover, our model is specifically devised to retain the key features associated with the anisotropy in the diffusion effects induced by the fiber architecture, with rapid variations across the thickness which cannot be adequately represented by naive averaging strategies. Our proposed model relies on a detailed asymptotic analysis in which we identify a limit model and establish strong convergence results. We also provide detailed numerical assessments which confirm an excellent accuracy of the surface-based model – compared with the reference 3D model – including in the representation of a complex phenomenon, namely, spiral waves.
- In [45] we assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model - derived and substantiated by mathematical arguments - is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy.

6.4. Lung and respiration modeling

Participants: Grégory Arbia, Laurent Boudin, Muriel Boulakia, Benoit Fabrèges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Stéphane Liwarek, Jessica Oakes, Ayman Moussa, Irène Vignon-Clementel.

- In [66] we are interested in the mathematical modeling of the deformation of the human lung tissue, called the lung parenchyma, during the respiration process. The parenchyma is a foam-like elastic material containing millions of air-filled alveoli connected by a tree- shaped network of airways. In this study, the parenchyma is governed by the linearized elasticity equations and the air movement in the tree by the Poiseuille law in each airway. The geometric arrangement of the alveoli is assumed to be periodic with a small period. We use the two-scale convergence theory to study the asymptotic behavior as the period goes to zero. The effect of the network of airways is described by a nonlocal operator and we propose a simple geometrical setting for which we show that this operator converges. We identify in the limit the equations modeling the homogenized behavior under an abstract convergence condition on this nonlocal operator. We derive some mechanical properties

of the limit material by studying the homogenized equations: the limit model is nonlocal both in space and time if the parenchyma material is considered compressible, but only in space if it is incompressible. Finally, we propose a numerical method to solve the homogenized equations and we study numerically a few properties of the homogenized parenchyma model.

- In [31] we present a calculation of the functioning of an acinus at exercise. We show that, given the geometry and the breathing dynamics of real acini, respiration can be correlated to a single equivalent parameter that we call the integrative permeability. We find that both V_{O2max} and PAO2 depend on this permeability in a non-linear manner.
- In [19], In this paper, we consider the Stokes equations and we are concerned with the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of the boundary. We first study the identifiability of the Robin coefficient and then we establish a stability estimate of logarithm type thanks to a Carleman inequality due to A. L. Bukhgeim and under the assumption that the velocity of a given reference solution stays far from 0 on a part of the boundary where Robin conditions are prescribed.
- In [18],In this work, we investigate the asymptotic behaviour of the solutions to the non-reactive fully elastic Boltzmann equations for mixtures in the diffusive scaling. We deal with cross sections such as hard spheres or cut-off power law potentials. We use Hilbert expansions near the common thermodynamic equilibrium granted by the H-theorem. The lower-order non trivial equality obtained from the Boltzmann equations leads to a linear functional equation in the velocity variable which is solved thanks to the Fredholm alternative. Since we consider multicomponent mixtures, the classical techniques introduced by Grad cannot be applied, and we propose a new method to treat the terms involving particles with different masses. The next-order equality in the Hilbert expansion then allows to write the macroscopic continuity equations for each component of the mixture.
- In [39], In this paper we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [40] we solved for the airflow and aerosol particle distribution in healthy and emphysematous rat lungs. Following our preliminary work in [79], we first estimated the respiratory resistance and compliance parameters from pressure measurements taken during ventilation experiments performed in healthy and emphysematous rats. Next, the 3D Navier-Stokes equations were solved in a Magnetic Resonance derived airway geometry coupled to 0D models at the boundaries leading to the five rat lobes. The multiscale 3D-0D simulations enabled consistent pressure and airflow results, unlike what was found when a constant pressure was described at the boundaries. Aerosolized particles were tracked throughout inspiration and the effects of particle size and gravity were studied. Healthy, homogeneous and heterogeneous disease cases were assessed. Once available, these in-silico predictions may be compared to experimental deposition data.

6.5. Miscellaneous

Participants: Grégory Arbia, Laurent Boudin, Jean-Frédéric Gerbeau, Damiano Lombardi, Marina Vidrascu, Irène Vignon-Clementel.

- In [13] we analyse the solution of the linear advection equation on a uniform mesh by a non dissipative second order scheme for discontinuous initial condition. We focus on the case of advection of a step function by the leapfrog scheme. We derive closed form exact and approximate solutions for the scheme that accurately predict oscillations of the numerical scheme.
- In [41] The recent biomechanical theory of cancer growth considers solid tumors as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumor into a host tissue is mainly driven by either the cell diffusion constant or the cell division rate, the latter depending either on the local cell density (contact inhibition), on mechanical stress in the tumor, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modeling, we prove there are always traveling waves above a minimal speed and we analyse their shapes. They appear to be complex with composite shapes and discontinuities. Several small parameters allow for analytical solutions; in particular the incompressible cells limit is very singular and related to the Hele-Shaw equation. These singular traveling waves are recovered numerically.
- In [68] This paper is devoted to the use of the entropy and duality methods for the existence theory of reaction-cross diffusion systems consisting of two equations, in any dimension of space. Those systems appear in population dynamics when the diffusion rates of individuals of two species depend on the concentration of individuals of the same species (self-diffusion), or of the other species (cross diffusion).
- In [65] We consider in this paper a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.
- In [42],We propose a method of modelling sail type structures which captures the wrinkling behaviour of such structures. The method is validated through experimental and analytical test cases, particularly in terms of wrinkling prediction. An enhanced wrinkling index is proposed as a valuable measure characterizing the global wrinkling development on the deformed structure. The method is based on a pseudo-dynamic finite element procedure involving non-linear MITC shell elements. The major advantage compared to membrane models generally used for this type of analysis is that no ad hoc wrinkling model is required to control the stability of the structure. We demonstrate our approach to analyse the behaviour of various structures with spherical and cylindrical shapes, characteristic of downwind sails over a rather wide range of shape and constitutive parameters. In all cases convergence is reached and the overall flying shape is most adequately represented, which shows that our approach is a most valuable alternative to standard techniques to provide deeper insight into the physical behaviour. Limitations appear only in some very special instances in which local wrinkling-related instabilities are extremely high and would require specific additional treatments, out of the scope of the present study.
- In [34],Since the pioneering work by Treloar, many models based on polymer chain statistics have been proposed to describe rubber elasticity. Recently, Alicandro, Cicalese, and the first author have rigorously derived a continuum theory of rubber elasticity from a discrete model by variational convergence. The aim of this paper is twofold. First we further physically motivate this model, and complete the analysis by numerical simulations. Second, in order to compare this model to the literature, we present in a common language two other representative types of models, specify their underlying assumptions, check their mathematical properties, and compare them to Treloar's experiments.
- In [48] We apply the domain decomposition method to linear elasticity problems for multi-materials

where the heterogeneities are concentrated in a thin internal layer. In the first case the heterogeneities are small, identical and periodically distributed on an internal surface and in the second one all the thin, curved internal layer is made of an elastic material much more strong than the surrounding one. In the first case the domain decomposition is used to efficiently solve the non-standard transmission problems obtained by the asymptotic expansion method. In the second case a non-standard membrane transmission problem originates from a surface shell like energy.

- In [32] : Our aim is to numerically validate the effectiveness of a matched asymptotic expansion formal method introduced in a pioneering paper by Nguetseng and Sànchez Palencia and extended in [76], [33]. Using this method a simplified model for the influence of small identical heterogeneities periodically distributed on an internal surface to the overall response of a linearly elastic body is derived. In order to validate this formal method a careful numerical study compares the solution obtained by a standard method on a fine mesh to the one obtained by asymptotic expansion. We compute both the zero and the first order terms in the expansion. To efficiently compute the first order term we introduce a suitable domain decomposition method.
- In [39] we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [65] we consider a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.

SECRET Project-Team

5. New Results

5.1. Symmetric cryptosystems

Participants: Anne Canteaut, Pascale Charpin, Virginie Lallemand, Gaëtan Leurent, María Naya-Plasencia, Joëlle Roué, Valentin Suder.

From outside, it might appear that symmetric techniques become obsolete after the invention of public-key cryptography in the mid 1970's. However, they are still widely used because they are the only ones that can achieve some major features like high-speed or low-cost encryption, fast authentication, and efficient hashing. Today, we find symmetric algorithms in GSM mobile phones, in credit cards, in WLAN connections. Symmetric cryptology is a very active research area which is stimulated by a pressing industrial demand for low-cost implementations (in terms of power consumption, gate complexity...). These extremely restricted implementation requirements are crucial when designing secure symmetric primitives and they might be at the origin of some weaknesses. Actually, these constraints seem quite incompatible with the rather complex mathematical tools needed for constructing a provably secure system.

The specificity of our research work is that it considers all aspects of the field, from the practical ones (new attacks, concrete specifications of new systems) to the most theoretical ones (study of the algebraic structure of underlying mathematical objects, definition of optimal objects). But, our purpose is to study these aspects not separately but as several sides of the same domain. Our approach mainly relies on the idea that, in order to guarantee a provable resistance to the known attacks and to achieve extremely good performance, a symmetric cipher must use very particular building blocks, whose algebraic structures may introduce unintended weaknesses. Our research work captures this conflict for all families of symmetric ciphers. It includes new attacks and the search for new building blocks which ensure both a high resistance to the known attacks and a low implementation cost. This work, which combines cryptanalysis and the theoretical study of discrete mathematical objects, is essential to progress in the formal analysis of the security of symmetric systems.

In this context, the very important challenges are the designs of low-cost ciphers and of secure hash functions. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimisation of the performance) of such primitives.

5.1.1. Hash functions

Following the recent attacks against almost all existing hash functions (MD5, SHA-0, SHA-1...), we have initiated a research work in this area, especially within the Saphir-2 ANR project and with several PhD theses. Our work on hash functions is two-fold: we have designed two new hash functions, named FSB and Shabal, which have been submitted to the SHA-3 competition, and we have investigated the security of several hash functions, including the new SHA-3 standard.

Recent results:

- Upper bounds on the degree of an iterated permutation from the degree of the inverse of the inner transformation; this result has been applied both to hash functions and to block ciphers. Most notably, this work leads to the best (theoretical) analysis of the hash function Keccak, which has been selected for the new SHA-3 standard [11].
- Study of a new technique for attacking symmetric primitives based on the existence of linear relations between some input and output bits of the Sbox. This method has been used for improving the best known attack against the SHA-3 candidate Hamsi [36], [58].

5.1.2. Block ciphers

Even if the security of the current block cipher standard, AES, is not threatened when it is used in a classical context, there is still a need for the design of improved attacks, and for the determination of design criteria which guarantee that the existing attacks do not apply. This notably requires a deep understanding of all previously proposed attacks. Moreover, there is a high demand from the industry of lightweight block ciphers for some constrained environments. Several such algorithms have been proposed in the last few years and their security should be carefully analysed. Most of our work in this area is related to an ANR Project named BLOC.

Recent results:

- Cryptanalysis of several recently proposed lightweight block ciphers. This includes an attack against the full cipher KLEIN-64 [66], [49], and an attack against 8 rounds (out of 12) of PRINCE [37].
- Analysis of the resistance of AES-like permutations to improved rebound attacks. Most notably, this improved technique leads to a distinguisher on 10 rounds of the internal permutation of the SHA-3 candidate Grøstl [14].
- Proposal of a new family of distinguishers against AES-based permutations, named *limited-birthday distinguishers*; these distinguishers exploit some some improved rebound techniques. They have been successfully applied to various AES-based primitives including AES, ECHO, Grøstl, LED, PHOTON and Whirlpool [42].
- Design of an improved variant of Meet-in-the-Middle attacks, named *Sieve-in-the-Middle*: instead of selecting the key candidates by searching for a collision in an intermediate state which can be computed forwards and backwards, we here look for the existence of valid transitions through some middle Sbox. In the same paper, an improved technique is also proposed to build bicliques without needing any additional data (on the contrary to classical biclique attacks). These new methods have been exploited to break 8 rounds (out of 12) of the lightweight block cipher PRINCE [37], [59], [30].
- Analysis of the differential properties of the AES Superbox [48].
- Design of a new block cipher, named ZORRO, for which physical security is considered as an optimisation criterion [41].
- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalises the so-called α-reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [24].

5.1.3. 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, characterising 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. For instance, bent functions, which are the Boolean functions which achieve the highest possible nonlinearity, have been extensively studied in order to provide some elements for a classification, or to adapt these functions to practical cryptographic constructions. We have also been interested in functions with a low differential uniformity (*e.g.*, APN functions), which are the S-boxes ensuring an (almost) optimal resistance to differential cryptanalysis.

Recent results:

- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [16], [51].
- Definition of a new criterion for Sboxes and link with some recent algebraic attacks on the hash function Hamsi [36], [58].
- Definition of some extended criterion for estimating the resistance of a block cipher to differential attacks. Most notably, this new criterion points out the fact that affinely equivalent Sboxes may not provide the same security level regarding differential cryptanalysis. 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 [21], [48].
- A new sufficient (and simpler) condition for checking that a mapping is APN has been established [62].
- Surveys of PN and APN mappings [55], [54].

5.2. Code-based cryptography

Participants: Grégory Landais, Rafael Misoczki, Nicolas Sendrier, Dimitrios Simos, Jean-Pierre Tillich.

Most popular public-key cryptographic schemes rely either on the factorisation problem (RSA, Rabin), or on the discrete logarithm problem (Diffie-Hellman, El Gamal, DSA). These systems have evolved and today instead of the classical groups ($\mathbf{Z}/n\mathbf{Z}$) we may use groups on elliptic curves. They allow a shorter block and key size for the same level of security. An intensive effort of the research community has been and is still being conducted to investigate the main aspects of these systems: implementation, theoretical and practical security. It must be noted that these systems all rely on algorithmic number theory. As they are used in most, if not all, applications of public-key cryptography today (and it will probably remain so in the near future), cryptographic applications are thus vulnerable to a single breakthrough in algorithmics or in hardware (a quantum computer can break all those schemes).

Diversity is a way to dilute that risk, and it is the duty of the cryptographic research community to prepare and propose alternatives to the number theoretic based systems. The most serious tracks today are lattice-based cryptography (NTRU,...), multivariate cryptography (HFE,...) and code-based cryptography (McEliece encryption scheme,...). All these alternatives are referred to as *post-quantum cryptosystems*, since they rely on difficult algorithmic problems which would not be solved by the coming-up of the quantum computer.

The code-based primitives have been investigated in details within the project-team. 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:

- Design of a new variant of McEliece using Moderate Density Parity Check (MDPC) codes [45];
- Cryptanalysis of McEliece system based on Wild Goppa codes from a quadratic finite field extension. This polynomial-time structural attack relies on some filtration of nested subcodes which will reveal the secret algebraic description of the underlying secret code [39], [63].
- Cryptanalysis of a variant of the McEliece cryptosystem based on Reed-Solomon codes [38].

- Cryptanalysis of a variant of the McEliece cryptosystem based on convolutional codes proposed by Löndahl and Johansson in 2012 [43].
- Design of the first algorithm for distinguishing between Goppa codes (or alternant codes) over any field and random codes. Provided that the codes have sufficiently large rates, this technique can solve in polynomial-time the Goppa-Code-Distinguishing problem, which is an assumption in the security proof of McEliece cryptosystem [12].
- Study of the hardness of the code equivalence problem over $\mathbf{F}_q$. This problem has been extensively studied for permutation-equivalence (which covers all cases for $q = 2$). For $q \in \{3, 4\}$, we have generalised the support-splitting algorithm, and we have shown that the problem seems intractable for most instances when $q \geq 5$ [46]. This property has been exploited in an improvement version of an identification protocol due to Girault [47].

5.3. Reverse engineering of communication systems

Participants: Marion Bellard, Nicolas Sendrier, Jean-Pierre Tillich, Audrey Tixier.

To assess the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle [1], it would be dangerous to rely, even partially, on the fact that the adversary does not know those specifications. However, this fundamental rule does not mean that the specifications are known to the attacker. In practice, before mounting a cryptanalysis, it is necessary to strip off the data. This reverse engineering process is often subtle, even when the data formatting is not concealed on purpose. A typical case is interception; some raw data, not necessarily encrypted, are observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. 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 DGA.

Recent results:

- Reconstruction of the constellation labelling (i.e. used in the modulator of a communication system) in the presence of errors and when the underlying code is convolutional (Marion Bellard's PhD).

5.4. Quantum information theory

Participants: André Chailloux, Anthony Leverrier, Denise Maurice, Jean-Pierre Tillich.

The field of Quantum Information and Computation aims at exploiting the laws of quantum physics to manipulate information in radically novel ways. Two main applications come to mind: quantum computers, that offer the promise of solving some problems intractable with classical computers (for instance, factorization); and quantum cryptography, which provides new ways to exchange data in a provably secure fashion.

The main obstacle towards the development of quantum computing is decoherence, a consequence of the interaction of the computer with a noisy environment. We investigate approaches to quantum error-correction as a way to fight against this effect, and we study more particularly some families of quantum error-correcting codes which generalise the best classical codes available today.

Our research also covers quantum cryptography where we study the security of efficient protocols for key distribution, in collaboration with experimental groups. More generally, we investigate how quantum theory severely constrains the action of honest and malicious parties in cryptographic scenarios.

5.4.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It 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. Our approach for overcoming this problem has been to study whether or not the family of turbo-codes and LDPC codes (and the associated iterative decoding algorithms) have a quantum counterpart.

[1] Kerckhoffs stated that principle in a paper entitled *La Cryptographie militaire*, published in 1883.

Recent results:

- Construction of quantum codes combining an improved version of a family of spatially coupled quantum LDPC codes with a family of error reducing turbo-codes [44];
- construction of quantum LDPC codes with fixed non-zero rate and a minimum distance which grows proportionally to the square root of the block-length. This greatly improves the previously best known construction whose minimum distance was logarithmic in the block-length [19].
- Mamdouh Abbara's PhD thesis [9]

5.4.2. 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 with security properties based on quantum theory.

Recent results:

- Experimental demonstration of quantum key distribution with continuous variables over 80 km [15], greatly improving over previous records around 25 km.
- Security proof of continuous-variable quantum key distribution protocols against general attacks [17], [29].
- Security proof of device-independent quantum key distribution in the bounded storage model [18].
- Study of BosonSampling, a recently introduced problem where quantum computers offer a provable speedup over classical computers [67], [28].
- Introduction and study of "Local Orthogonality", an information-theoretical principle for quantum correlations [13], [68].
- Introduction of a general formalism for the study of contextuality and non locality in quantum theory, based on the combinatorics of hypergraphs [65], [27].

SIERRA Project-Team

6. New Results

6.1. Block-Coordinate Frank-Wolfe Optimization for Structural SVMs

Participants: Simon Lacoste-Julien [correspondent], Mark Schmidt.

Collaboration with: Martin Jaggi (Centre de Mathématiques Appliquées, Ecole Polytechnique), Patrick Pletscher (Machine Learning Laboratory, ETH Zurich).

In [16] we propose a randomized block-coordinate variant of the classic Frank-Wolfe algorithm for convex optimization with block-separable constraints. Despite its lower iteration cost, we show that it achieves a similar convergence rate in duality gap as the full Frank-Wolfe algorithm. We also show that, when applied to the dual structural support vector machine (SVM) objective, it yields an online algorithm that has the same low iteration complexity as primal stochastic subgradient methods. However, unlike stochastic subgradient methods, the block-coordinate Frank-Wolfe algorithm allows us to compute the optimal step-size and yields a computable duality gap guarantee. Our experiments indicate that this simple algorithm outperforms competing structural SVM solvers.

6.2. Minimizing Finite Sums with the Stochastic Average Gradient.

Participants: Mark Schmidt [correspondent], Nicolas Le Roux, Francis Bach.

In [32] we propose the stochastic average gradient (SAG) method for optimizing the sum of a finite number of smooth convex functions. Like stochastic gradient (SG) methods, the SAG method's iteration cost is independent of the number of terms in the sum. However, by incorporating a memory of previous gradient values the SAG method achieves a faster convergence rate than black-box SG methods. The convergence rate is improved from $O(1/\sqrt{k})$ to $O(1/k)$ in general, and when the sum is strongly-convex the convergence rate is improved from the sub-linear $O(1/k)$ to a linear convergence rate of the form $O(\rho^k)$ for $\rho < 1$. Further, in many cases the convergence rate of the new method is also faster than black-box deterministic gradient methods, in terms of the number of gradient evaluations. Numerical experiments indicate that the new algorithm often dramatically outperforms existing SG and deterministic gradient methods.

The primary contribution of this work is the analysis of a new algorithm that we call the *stochastic average gradient* (SAG) method, a randomized variant of the incremental aggregated gradient (IAG) method of [43]. The SAG method has the low iteration cost of SG methods, but achieves the convergence rates stated above for the FG method. The SAG iterations take the form

$$x^{k+1} = x^k - \frac{\alpha_k}{n} \sum_{i=1}^{n} y_i^k, \tag{1}$$

where at each iteration a random index i_k is selected and we set $y_i^k = f_i'(x^k)$ if $i = i_k$, and y_i^{k-1} otherwise. That is, like the FG method, the step incorporates a gradient with respect to each function. But, like the SG method, each iteration only computes the gradient with respect to a single example and the cost of the iterations is independent of n. Despite the low cost of the SAG iterations, we show in this paper that with a constant step-size *the SAG iterations have an $O(1/k)$ convergence rate for convex objectives and a linear convergence rate for strongly-convex objectives*, like the FG method. That is, by having access to i_k and by keeping a *memory* of the most recent gradient value computed for each index i, this iteration achieves a faster convergence rate than is possible for standard SG methods. Further, in terms of effective passes through the data, we will also see that for many problems the convergence rate of the SAG method is also faster than is possible for standard FG methods.

6.3. Fast Convergence of Stochastic Gradient Descent under a Strong Growth Condition

Participants: Mark Schmidt [correspondent], Nicolas Le Roux [correspondent].

In [33] we consider optimizing a function smooth convex function f that is the average of a set of differentiable functions f_i, under the assumption considered by [87] and [90] that the norm of each gradient f_i' is bounded by a linear function of the norm of the average gradient f'. We show that under these assumptions the basic stochastic gradient method with a sufficiently-small constant step-size has an $O(1/k)$ convergence rate, and has a linear convergence rate if g is strongly-convex.

We write our problem

$$\min_{x \in \mathbb{R}^P} f(x) := \frac{1}{N} \sum_{i=1}^{N} f_i(x), \tag{2}$$

where we assume that f is convex and its gradient f' is Lipschitz-continuous with constant L, meaning that for all x and y we have

$$||f'(x) - f'(y)|| \leq L||x - y||.$$

If f is twice-differentiable, these assumptions are equivalent to assuming that the eigenvalues of the Hessian $f''(x)$ are bounded between 0 and L for all x.

Deterministic gradient methods for problems of this form use the iteration

$$x_{k+1} = x_k - \alpha_k f'(x_k), \tag{3}$$

for a sequence of step sizes α_k. In contrast, *stochastic gradient* methods use the iteration

$$x_{k+1} = x_k - \alpha_k f_i'(x_k), \tag{4}$$

for an individual data sample i selected uniformly at random from the set $\{1, 2, \cdots, N\}$.

The stochastic gradient method is appealing because the cost of its iterations is *independent of*N. However, in order to guarantee convergence stochastic gradient methods require a decreasing sequence of step sizes $\{\alpha_k\}$ and this leads to a slower convergence rate. In particular, for convex objective functions the stochastic gradient method with a decreasing sequence of step sizes has an expected error on iteration k of $O(1/\sqrt{k})$ [78], meaning that

$$\mathbb{E}[f(x_k)] - f(x^*) = O(1/\sqrt{k}).$$

In contrast, the deterministic gradient method with a *constant* step size has a smaller error of $O(1/k)$ [79]. The situation is more dramatic when f is *strongly* convex, meaning that

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle + \frac{\mu}{2}||y - x||^2, \tag{5}$$

for all x and y and some $\mu > 0$. For twice-differentiable functions, this is equivalent to assuming that the eigenvalues of the Hessian are bounded below by μ. For strongly convex objective functions, the stochastic gradient method with a decreasing sequence of step sizes has an error of $O(1/k)$ [77] while the deterministic method with a constant step size has an *linear* convergence rate. In particular, the deterministic method satisfies

$$f(x_k) - f(x^*) \leq \rho^k [f(x_0) - f(x^*)],$$

for some $\rho < 1$ [71].

We show that if the individual gradients $f_i'(x_k)$ satisfy a certain strong growth condition relative to the full gradient $f'(x_k)$, the stochastic gradient method with a sufficiently small constant step size achieves (in expectation) the convergence rates stated above for the deterministic gradient method.

6.4. Non-strongly-convex smooth stochastic approximation with convergence rate $O(1/n)$

Participants: Eric Moulines, Francis Bach [correspondent].

Large-scale machine learning problems are becoming ubiquitous in many areas of science and engineering. Faced with large amounts of data, practitioners typically prefer algorithms that process each observation only once, or a few times. Stochastic approximation algorithms such as stochastic gradient descent (SGD) and its variants, although introduced more than sixty years ago, still remain the most widely used and studied method in this context. In [8], we consider the stochastic approximation problem where a convex function has to be minimized, given only the knowledge of unbiased estimates of its gradients at certain points, a framework which includes machine learning methods based on the minimization of the empirical risk. We focus on problems without strong convexity, for which all previously known algorithms achieve a convergence rate for function values of $O(1/\sqrt{n})$ after n iterations. We consider and analyze two algorithms that achieve a rate of $O(1/n)$ for classical supervised learning problems. For least-squares regression, we show that *averaged* stochastic gradient descent *with constant step-size* achieves the desired rate. For logistic regression, this is achieved by a simple novel stochastic gradient algorithm that (a) constructs successive local quadratic approximations of the loss functions, while (b) preserving the same running-time complexity as stochastic gradient descent. For these algorithms, we provide a non-asymptotic analysis of the generalization error (in expectation, and also in high probability for least-squares), and run extensive experiments showing that they often outperform existing approaches.

6.5. Streaming Bayesian Inference

Participant: Michael Jordan [correspondent].

Large, streaming data sets are increasingly the norm in science and technology. Simple descriptive statistics can often be readily computed with a constant number of operations for each data point in the streaming setting, without the need to revisit past data or have advance knowledge of future data. But these time and memory restrictions are not generally available for the complex, hierarchical models that practitioners often have in mind when they collect large data sets. Significant progress on scalable learning procedures has been made in recent years. But the underlying models remain simple, and the inferential framework is generally non-Bayesian. The advantages of the Bayesian paradigm (e.g., hierarchical modeling, coherent treatment of uncertainty) currently seem out of reach in the Big Data setting.

An exception to this statement is provided by Hofmann et al. (2010), who have shown that a class of approximation methods known as *variational Bayes* (VB) can be usefully deployed for large-scale data sets. They have applied their approach, referred to as *stochastic variational inference* (SVI), to the domain of topic modeling of document collections, an area with a major need for scalable inference algorithms. VB traditionally uses the variational lower bound on the marginal likelihood as an objective function, and the idea of SVI is to apply a variant of stochastic gradient descent to this objective. Notably, this objective is based on the conceptual existence of a full data set involving D data points (i.e., documents in the topic model setting), for a fixed value of D. Although the stochastic gradient is computed for a single, small subset of data points (documents) at a time, the posterior being targeted is a posterior for D data points. This value of D must be specified in advance and is used by the algorithm at each step. Posteriors for D' data points, for D' not equal to D, are not obtained as part of the analysis.

We view this lack of a link between the number of documents that have been processed thus far and the posterior that is being targeted as undesirable in many settings involving streaming data. In this project we aim at an approximate Bayesian inference algorithm that is scalable like SVI but is also truly a streaming procedure, in that it yields an approximate posterior for each processed collection of D' data points—and not just a pre-specified "final" number of data points D. To that end, we return to the classical perspective of Bayesian updating, where the recursive application of Bayes theorem provides a sequence of posteriors, not a sequence of approximations to a fixed posterior. To this classical recursive perspective we bring the VB framework; our updates need not be exact Bayesian updates but rather may be approximations such as VB.

Although the empirical success of SVI is the main motivation for our work, we are also motivated by recent developments in computer architectures, which permit distributed and asynchronous computations in addition to streaming computations. A streaming VB algorithm naturally lends itself to distributed and asynchronous implementations.

6.6. Convex Relaxations for Permutation Problems

Participants: Fajwel Fogel [correspondent], Rodolphe Jenatton, Francis Bach, Alexandre d'Aspremont.

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 [12] 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 performed numerical experiments on archeological data, Markov chains and gene sequences.

6.7. Phase retrieval for imaging problems

Participants: Fajwel Fogel [correspondent], Irène Waldspurger, Alexandre d'Aspremont.

In [29] we study convex relaxation algorithms for phase retrieval on imaging problems. We show that structural assumptions on the signal and the observations, such as sparsity, smoothness or positivity, can be exploited to both speed-up convergence and improve recovery performance. We detail experimental results in molecular imaging problems simulated from PDB data.

Phase retrieval seeks to reconstruct a complex signal, given a number of observations on the *magnitude* of linear measurements, i.e. solve

$$\begin{array}{ll} \text{find} & x \\ \text{such that} & |Ax| = b \end{array}$$

in the variable x, where A and b. This problem has direct applications in X-ray and crystallography imaging, diffraction imaging, Fourier optics or microscopy for example, in problems where physical limitations mean detectors usually capture the intensity of observations but cannot recover their phase. In this project, we focus on problems arising in diffraction imaging, where A is usually a Fourier transform, often composed with one or multiple masks (a technique sometimes called ptychography). The Fourier structure, through the FFT, often considerably speeds up basic linear operations, which allows us to solve large scale convex relaxations on realistically large imaging problems. We also observe that in most of the imaging problems we consider, the Fourier transform is very sparse, with known support (we lose the phase but observe the magnitude of Fourier coefficients), which allows us to considerably reduce the size of our convex phase retrieval relaxations.

6.8. Learning Sparse Penalties for Change-point Detection using Max Margin Interval Regression

Participants: Toby Hocking, Guillem Rigaill, Jean-Philippe Vert, Francis Bach [correspondent].

In segmentation models, the number of change-points is typically chosen using a penalized cost function. In [22] we propose to learn the penalty and its constants in databases of signals with weak change-point annotations. We propose a convex relaxation for the resulting interval regression problem, and solve it using accelerated proximal gradient methods. We show that this method achieves state-of-the-art change-point detection in a database of annotated DNA copy number profiles from neuroblastoma tumors.

6.9. Maximizing submodular functions using probabilistic graphical models

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

In [34] we consider the problem of maximizing submodular functions; while this problem is known to be NP-hard, several numerically efficient local search techniques with approximation guarantees are available. In this paper, we propose a novel convex relaxation which is based on the relationship between submodular functions, entropies and probabilistic graphical models. In a graphical model, the entropy of the joint distribution decomposes as a sum of marginal entropies of subsets of variables; moreover, for any distribution, the entropy of the closest distribution factorizing in the graphical model provides an bound on the entropy. For directed graphical models, this last property turns out to be a direct consequence of the submodularity of the entropy function, and allows the generalization of graphical-model-based upper bounds to any submodular functions. These upper bounds may then be jointly maximized with respect to a set, while minimized with respect to the graph, leading to a convex variational inference scheme for maximizing submodular functions, based on outer approximations of the marginal polytope and maximum likelihood bounded treewidth structures. By considering graphs of increasing treewidths, we may then explore the trade-off between computational complexity and tightness of the relaxation. We also present extensions to constrained problems and maximizing the difference of submodular functions, which include all possible set functions.

Optimizing submodular functions has been an active area of research with applications in graph-cut-based image segmentation [44], sensor placement [69], or document summarization [70]. A set function F is a function defined on the power set 2^V of a certain set V. It is submodular if and only if for all $A, B \subseteq V$, $F(A) + F(B) \geqslant F(A \cap B) + F(A \cup B)$. Equivalently, these functions also admit the diminishing returns property, i.e., the marginal cost of an element in the context of a smaller set is more than its cost in the context of a larger set. Classical examples of such functions are entropy, mutual information, cut functions, and covering functions—see further examples in [58], [38].

Submodular functions form an interesting class of discrete functions because minimizing a submodular function can be done in polynomial time [58], while maximization, although NP-hard, admits constant factor approximation algorithms [76]. In this paper, our ultimate goal is to provide the first (to the best of our knowledge) generic convex relaxation of submodular function maximization, with a hierarchy of complexities related to known combinatorial hierarchies such as the Sherali-Adams hierarchy [83]. Beyond the graphical model tools that we are going to develop, having convex relaxations may be interesting for several reasons:

(1) they can lead to better solutions, (2) they provide online bounds that may be used within branch-and-bound optimization and (3) they ease the use of such combinatorial optimization problems within structured prediction framework [91].

We make the following contributions:

- For any directed acyclic graph G and a submodular function F, we define a bound $F_G(A)$ and study its properties (monotonicity, tightness), which is specialized to decomposable graphs.
- We propose an algorithm to maximize submodular functions by maximizing the bound $F_G(A)$ with respect to A while minimizing with respect to the graph G, leading to a convex variational method based on outer approximation of the marginal polytope [93] and inner approximation of the hypertree polytope.
- We propose extensions to constrained problems and maximizing the difference of submodular functions, which include all possible set functions.
- We illustrate our results on small-scale experiments.

6.10. Reflection methods for user-friendly submodular optimization

Participants: Stefanie Jegelka, Suvrit Sra, Francis Bach [correspondent].

Recently, it has become evident that submodularity naturally captures widely occurring concepts in machine learning, signal processing and computer vision. Consequently, there is need for efficient optimization procedures for submodular functions, especially for minimization problems. While general submodular minimization is challenging, we propose in [15] a new method that exploits existing decomposability of submodular functions. In contrast to previous approaches, our method is neither approximate, nor impractical, nor does it need any cumbersome parameter tuning. Moreover, it is easy to implement and parallelize. A key component of our method is a formulation of the discrete submodular minimization problem as a continuous best approximation problem that is solved through a sequence of reflections, and its solution can be easily thresholded to obtain an optimal discrete solution. This method solves *both* the continuous and discrete formulations of the problem, and therefore has applications in learning, inference, and reconstruction. In our experiments, we illustrate the benefits of our method on two image segmentation tasks.

6.11. Convex Relaxations for Learning Bounded Treewidth Decomposable Graphs

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

In [24] we consider the problem of learning the structure of undirected graphical models with bounded treewidth, within the maximum likelihood framework. This is an NP-hard problem and most approaches consider local search techniques. In this paper, we pose it as a combinatorial optimization problem, which is then relaxed to a convex optimization problem that involves searching over the forest and hyperforest polytopes with special structures, independently. A supergradient method is used to solve the dual problem, with a run-time complexity of $O(k^3 n^{k+2} \log n)$ for each iteration, where n is the number of variables and k is a bound on the treewidth. We compare our approach to state-of-the-art methods on synthetic datasets and classical benchmarks, showing the gains of the novel convex approach.

Graphical models provide a versatile set of tools for probabilistic modeling of large collections of interdependent variables. They are defined by graphs that encode the conditional independences among the random variables, together with potential functions or conditional probability distributions that encode the specific local interactions leading to globally well-defined probability distributions [42], [93], [67].

In many domains such as computer vision, natural language processing or bioinformatics, the structure of the graph follows naturally from the constraints of the problem at hand. In other situations, it might be desirable to estimate this structure from a set of observations. It allows (a) a statistical fit of rich probability distributions that can be considered for further use, and (b) discovery of structural relationship between different variables. In the former case, distributions with tractable inference are often desirable, i.e., inference with run-time complexity does not scale exponentially in the number of variables in the model. The simplest constraint to ensure tractability is to impose tree-structured graphs [52]. However, these distributions are not rich enough, and following earlier work [73], [39], [75], [48], [59], [89], we consider models with *treewidth* bounded, not simply by one (i.e., trees), but by a small constant k.

Beyond the possibility of fitting tractable distributions (for which probabilistic inference has linear complexity in the number of variables), learning bounded-treewidth graphical models is key to design approximate inference algorithms for graphs with higher treewidth. Indeed, as shown by [82], [93], [68], approximating general distributions by tractable distributions is a common tool in variational inference. However, in practice, the complexity of variational distributions is often limited to trees (i.e., $k = 1$), since these are the only ones with exact polynomial-time structure learning algorithms. The convex relaxation we designed enables us to augment the applicability of variational inference, by allowing a finer trade-off between run-time complexity and approximation quality.

We make the following contributions:

- We provide a novel convex relaxation for learning bounded-treewidth decomposable graphical models from data in polynomial time. This is achieved by posing the problem as a combinatorial optimization problem, which is relaxed to a convex optimization problem that involves the graphic and hypergraphic matroids.
- We show how a supergradient ascent method may be used to solve the dual optimization problem, using greedy algorithms as inner loops on the two matroids. Each iteration has a run-time complexity of $O(k^3 n^{k+2} \log n)$, where n is the number of variables. We also show how to round the obtained fractional solution.
- We compare our approach to state-of-the-art methods on synthetic datasets and classical benchmarks showing the gains of the novel convex approach.

6.12. Large-Margin Metric Learning for Partitioning Problems

Participants: Rémi Lajugie [correspondent], Sylvain Arlot, Francis Bach.

In [31] we consider unsupervised partitioning problems, such as clustering, image segmentation, video segmentation and other change-point detection problems. We focus on partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, which include mean-based change-point detection, K-means, spectral clustering and normalized cuts. Our main goal is to learn a Mahalanobis metric for these unsupervised problems, leading to feature weighting and/or selection. This is done in a supervised way by assuming the availability of several potentially partially labelled datasets that share the same metric. We cast the metric learning problem as a large-margin structured prediction problem, with proper definition of regularizers and losses, leading to a convex optimization problem which can be solved efficiently with iterative techniques. We provide experiments where we show how learning the metric may significantly improve the partitioning performance in synthetic examples, bioinformatics, video segmentation and image segmentation problems.

Unsupervised partitioning problems are ubiquitous in machine learning and other data-oriented fields such as computer vision, bioinformatics or signal processing. They include (a) traditional *unsupervised clustering* problems, with the classical K-means algorithm, hierarchical linkage methods [61] and spectral clustering [80], (b) *unsupervised image segmentation* problems where two neighboring pixels are encouraged to be in the same cluster, with mean-shift techniques [51] or normalized cuts [84], and (c) *change-point detection* problems adapted to multivariate sequences (such as video) where segments are composed of contiguous elements, with typical window-based algorithms [54] and various methods looking for a change in the mean of the features (see, e.g., [49]).

All the algorithms mentioned above rely on a specific distance (or more generally a similarity measure) on the space of configurations. A good metric is crucial to the performance of these partitioning algorithms and its choice is heavily problem-dependent. While the choice of such a metric has been originally tackled manually (often by trial and error), recent work has considered learning such metric directly from data. Without any supervision, the problem is ill-posed and methods based on generative models may learn a metric or reduce dimensionality (see, e.g., [53]), but typically with no guarantees that they lead to better partitions. In this paper, we follow [41], [95], [40] and consider the goal of learning a metric for potentially several partitioning problems sharing the same metric, assuming that several fully or partially labelled partitioned datasets are available during the learning phase. While such labelled datasets are typically expensive to produce, there are several scenarios where these datasets have already been built, often for evaluation purposes. These occur in video segmentation tasks, image segmentation tasks as well as change-point detection tasks in bioinformatics (see [62]).

W consider partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, which include K-means, spectral clustering and normalized cuts, and mean-based change-point detection. We make the following contributions:

- We review and unify several partitioning algorithms, and cast them as the maximization of a linear function of a rescaled equivalence matrix, which can be solved by algorithms based on spectral relaxations or dynamic programming.
- Given fully labelled datasets, we cast the metric learning problem as a large-margin structured prediction problem, with proper definition of regularizers, losses and efficient loss-augmented inference.
- Given partially labelled datasets, we propose an algorithm, iterating between labeling the full datasets given a metric and learning a metric given the fully labelled datasets. We also consider extensions that allow changes in the full distribution of univariate time series (rather than changes only in the mean), with application to bioinformatics.
- We provide experiments where we show how learning the metric may significantly improve the partitioning performance in synthetic examples, video segmentation and image segmentation problems.

6.13. Comparison between multi-task and single-task oracle risks in kernel ridge regression

Participant: Matthieu Solnon [correspondent].

In [35] we study multi-task kernel ridge regression and try to understand when the multi-task procedure performs better than the single-task one, in terms of averaged quadratic risk. In order to do so, we compare the risks of the estimators with perfect calibration, the oracle risk. We are able to give explicit settings, favorable to the multi-task procedure, where the multi-task oracle performs better than the single-task one. In situations where the multi-task procedure is conjectured to perform badly, we also show the oracle does so. We then complete our study with simulated examples, where we can compare both oracle risks in more natural situations. A consequence of our result is that the multi-task ridge estimator has a lower risk than any single-task estimator, in favorable situations.

Increasing the sample size is the most common way to improve the performance of statistical estimators. In some cases (see, for instance, the experiments of [56] on customer data analysis or those of [63] on molecule binding problems), having access to some new data may be impossible, often due to experimental limitations. One way to circumvent those constraints is to use datasets from several related (and, hopefully, "similar") problems, as if it gave additional (in some sense) observations on the initial problem. The statistical methods using this heuristic are called "multi-task" techniques, as opposed to "single-task" techniques, where every problem is treated one at a time. In this paper, we study kernel ridge regression in a multi-task framework and try to understand when multi-task can improve over single-task.

The first trace of a multi-task estimator can be found in the work of [88]. In this article, Charles Stein showed that the usual maximum-likelihood estimator of the mean of a Gaussian vector (of dimension larger than 3, every dimension representing here a task) is not admissible—that is, there exists another estimator that has a lower risk for every parameter. He showed the existence of an estimator that uniformly attains a lower quadratic risk by shrinking the estimators along the different dimensions towards an arbitrary point. An explicit form of such an estimator was given by [64], yielding the famous James-Stein estimator. This phenomenon, now known as the "Stein's paradox", was widely studied in the following years and the behaviour of this estimator was confirmed by empirical studies, in particular the one from [55]. This first example clearly shows the goals of the multi-task procedure: an advantage is gained by borrowing information from different tasks (here, by shrinking the estimators along the different dimensions towards a common point), the improvement being scored by the global (averaged) squared risk. Therefore, this procedure does not guarantee individual gains on every task, but a global improvement on the sum of those task-wise risks.

We consider $p \geq 2$ different regression tasks, a framework we refer to as "multi-task" regression, and where the performance of the estimators is measured by the fixed-design quadratic risk. Kernel ridge regression is a classical framework to work with and comes with a natural norm, which often has desirable properties (such as, for instance, links with regularity). This norm is also a natural "similarity measure" between the regression functions. [56] showed how to extend kernel ridge regression to a multi-task setting, by adding a regularization term that binds the regression functions along the different tasks together. One of the main questions that is asked is to assert whether the multi-task estimator has a lower risk than any single-task estimator. It was recently proved by [86] that a fully data-driven calibration of this procedure is possible, given some assumptions on the set of matrices used to regularize—which correspond to prior knowledge on the tasks. Under those assumptions, the estimator is showed to verify an *oracle inequality*, that is, its risk matches (up to constants) the best possible one, the *oracle risk*. Thus, it suffices to compare the oracle risks for the multi-task procedure and the single-task one to provide an answer to this question.

We study the oracle multi-task risk and compare it to the oracle single-task risk. We then find situations where the multi-task oracle is proved to have a lower risk than the single-task oracle. This allows us to better understand which situation favors the multi-task procedure and which does not. After having defined our model, we write down the risk of a general multi-task ridge estimator and see that it admits a convenient decomposition using two key elements: the mean of the tasks and the resulting variance. This decomposition allows us to optimize this risk and get a precise estimation of the oracle risk, in settings where the ridge estimator is known to be minimax optimal. We then explore several repartitions of the tasks that give the latter multi-task rates, study their single-task oracle risk and compare it to their respective multi-task rates. This allows us to discriminate several situations, depending whether the multi-task oracle either outperforms its single-task counterpart, underperforms it or whether both behave similarly. We also show that, in the cases favorable to the multi-task oracle detailed in the previous sections, the estimator proposed by [86] behaves accordingly and achieves a lower risk than the single-task oracle. We finally study settings where we can no longer explicitly study the oracle risk, by running simulations, and we show that the multi-task oracle continues to retain the same virtues and disadvantages as before.

6.14. Sharp analysis of low-rank kernel matrix approximations

Participant: Francis Bach [correspondent].

Kernel methods, such as the support vector machine or kernel ridge regression, are now widely used in many areas of science and engineering, such as computer vision or bioinformatics. However, kernel methods typically suffer from at least quadratic running-time complexity in the number of observations n, as this is the complexity of computing the kernel matrix. In large-scale settings where n may be large, this is usually not acceptable. In [7], we consider supervised learning problems within the positive-definite kernel framework, such as kernel ridge regression, kernel logistic regression or the support vector machine. Low-rank approximations of the kernel matrix are often considered as they allow the reduction of running time complexities to $O(p^2 n)$, where p is the rank of the approximation. The practicality of such methods thus depends on the required rank p. In this paper, we show that in the context of kernel ridge regression, for

approximations based on a random subset of columns of the original kernel matrix, the rank p may be chosen to be linear in the *degrees of freedom* associated with the problem, a quantity which is classically used in the statistical analysis of such methods, and is often seen as the implicit number of parameters of non-parametric estimators. This result enables simple algorithms that have sub-quadratic running time complexity, but provably exhibit the same *predictive performance* than existing algorithms, for any given problem instance, and not only for worst-case situations.

6.15. fMRI encoding and decoding models

Participant: Fabian Pedregosa [correspondent].

In [20] we show that HRF estimation improves sensitivity of fMRI encoding and decoding models and propose a new approach for the estimation of Hemodynamic Response Functions from fMRI data. The model we propose is based on the linearity assumption behind the General Linear Model and can be computed using standard gradient-based solvers. We use the activation patterns computed by our model as input data for encoding and decoding studies and report performance improvement in both settings.

This work proves that significant improvements in recovery of brain activation patterns can be made by estimating the form of the Hemodynamic Response Function instead of using a canonical form for this response.

6.16. Structured Penalties for Log-linear Language Models

Participants: Anil Nelakanti [correspondent], Cédric Archambeau, Francis Bach, Guillaume Bouchard.

Language models can be formalized as log-linear regression models where the input features represent previously observed contexts up to a certain length m. The complexity of existing algorithms to learn the parameters by maximum likelihood scale linearly in nd, where n is the length of the training corpus and d is the number of observed features. In [19] we present a model that grows logarithmically in d, making it possible to efficiently leverage longer contexts. We account for the sequential structure of natural language using tree-structured penalized objectives to avoid overfitting and achieve better generalization.

Language models are crucial parts of advanced natural language processing pipelines, such as speech recognition [45], machine translation [47], or information retrieval [92]. When a sequence of symbols is observed, a language model predicts the probability of occurrence of the next symbol in the sequence. Models based on so-called back-off smoothing have shown good predictive power [60]. In particular, Kneser-Ney (KN) and its variants [66] are still achieving state-of-the-art results for more than a decade after they were originally proposed. Smoothing methods are in fact clever heuristics that require tuning parameters in an ad-hoc fashion. Hence, more principled ways of learning language models have been proposed based on maximum entropy [50] or conditional random fields [81], or by adopting a Bayesian approach [94].

We focus on penalized maximum likelihood estimation in log-linear models. In contrast to language models based on *unstructured* norms such as ℓ_2 (quadratic penalties) or ℓ_1 (absolute discounting), we use *tree-structured* norms [96], [65]. Structured penalties have been successfully applied to various NLP tasks, including chunking and named entity recognition [74], but not language modeling. Such penalties are particularly well-suited to this problem as they mimic the nested nature of word contexts. However, existing optimizing techniques are not scalable for large contexts m.

We show that structured tree norms provide an efficient framework for language modeling. Furthermore, we give the first algorithm for structured ℓ_∞ tree norms with a complexity nearly linear in the number of nodes. This leads to a memory-efficient *and* time-efficient learning algorithm for generalized linear language models.

6.17. Distributed Large-scale Natural Graph Factorization

Participants: Amr Ahmed, Nino Shervashidze [correspondent], Shravan Narayanamurthy, Vanja Josifovski, Alexander Smola.

Natural graphs, such as social networks, email graphs, or instant messaging patterns, have become pervasive through the internet. These graphs are massive, often containing hundreds of millions of nodes and billions of edges. While some theoretical models have been proposed to study such graphs, their analysis is still difficult due to the scale and nature of the data. We propose a framework for large-scale graph decomposition and inference. To resolve the scale, our framework in [6] is distributed so that the data are partitioned over a shared-nothing set of machines. We propose a novel factorization technique that relies on partitioning a graph so as to minimize the number of neighboring vertices rather than edges across partitions. Our decomposition is based on a streaming algorithm. It is network-aware as it adapts to the network topology of the underlying computational hardware. We use local copies of the variables and an efficient asynchronous communication protocol to synchronize the replicated values in order to perform most of the computation without having to incur the cost of network communication. On a graph of 200 million vertices and 10 billion edges, derived from an email communication network, our algorithm retains convergence properties while allowing for almost linear scalability in the number of computers.

6.18. Evaluating Speech Features with the Minimal-Pair ABX task

Participants: Thomas Schatz [correspondent], Vijayaditya Peddinti, Francis Bach, Aren Jansen, Hynek Hermansky, Emmanuel Dupoux.

In [23] we introduce a new framework for the evaluation of speech representations in zero-resource settings, that extends and complements previous work by Carlin, Jansen and Hermansky [46]. In particular, we replace their Same/Different discrimination task by several Minimal-Pair ABX (MP-ABX) tasks. We explain the analytical advantages of this new framework and apply it to decompose the standard signal processing pipelines for computing PLP and MFC coefficients. This method enables us to confirm and quantify a variety of well-known and not-so-well-known results in a single framework.

Speech recognition technology crucially rests on adequate speech features for encoding input data. Several such features have been proposed and studied (MFCCs, PLPs, etc), but they are often evaluated indirectly using complex tasks like phone classification or word identification. Such an evaluation technique suffers from several limitations. First, it requires a large enough annotated corpus in order to train the classifier/recognizer. Such a resource may not be available in all languages or dialects (the so-called "zero or limited resource" setting). Second, supervised classifiers may be too powerful and may compensate for potential defects of speech features (for instance noisy/unreliable channels). However, such defects are problematic in unsupervised learning techniques. Finally, the particular statistical assumptions of the classifier (linear, Gaussian, etc.) may not be suited for specific speech features (for instance sparse neural codes as in Hermansky [85]). It is therefore important to replace these complex evaluation schemes by simpler ones which tap more directly the properties of the speech features.

We extend and complement the framework proposed by Carlin, Jansen and Hermansky [46] for the evaluation of speech features in zero resource settings. This framework uses a Same-Different word discrimination task that does not depend on phonetically labelled data, nor on training a classifier. It assumes a speech corpus segmented into words, and derives a word-by-word acoustic distance matrix computed by comparing every word with every other one using Dynamic Time Warping (DTW). Carlin et al. then compute an average precision score which is used to evaluate speech features (the higher average precision, the better the features).

We explore an extension of this technique through Minimal-Pair ABX tasks (MP-ABX tasks) tested on a phonetically balanced corpus [57]. This improves the interpretability of the Carlin et al evaluation results in three different ways. First, the Same/Different task requires the computation of a ROC curve in order to derive average precision. In contrast, the ABX task is a discrimination task used in psychophysics (see [72], chapter 9) which allows for the direct computation of an error rate or a d' measure that are easier to interpret than average precision [46] and involve no assumption about ROC curves. Second, the Same/Different task compares *sets of words*, and as a result is influenced by the mix of similar versus distinct words or short versus long words in the corpus. The ABX task, in contrast, is computed on *word pairs*, and therefore enables to make linguistically precise comparisons, as in word *minimal pairs*, i.e. words differing by only one phoneme. Variants of the task enable to study phoneme discrimination across talkers and/or phonetic contexts, as well as

talker discrimination across phonemes. Because it is more controlled and provides a parameter and model-free metric, the MP-ABX error rate also enables to compare performance across databases or across languages. Third, we compute bootstrap-based estimates of the variability of our performance measures, which allows us to derive confidence intervals for the error rates and tests of the significance of the difference between the error rates obtained with different representations.

6.19. Hidden Markov Tree Models for Semantic Class Induction

Participants: Edouard Grave [correspondent], Guillaume Obozinski, Francis Bach.

In [13] we propose a new method for semantic class induction. First, we introduce a generative model of sentences, based on dependency trees and which takes into account homonymy. Our model can thus be seen as a generalization of Brown clustering. Second, we describe an efficient algorithm to perform inference and learning in this model. Third, we apply our proposed method on two large datasets (10^8 tokens, 10^5 words types), and demonstrate that classes induced by our algorithm improve performance over Brown clustering on the task of semi-supervised supersense tagging and named entity recognition.

Most competitive learning methods for computational linguistics are supervised, and thus require labeled examples, which are expensive to obtain. Moreover, those techniques suffer from data scarcity: many words only appear a small number of time, or even not at all, in the training data. It thus helps a lot to first learn word clusters on a large amount of unlabeled data, which are cheap to obtain, and then to use this clusters as features for the supervised task. This scheme has proven to be effective for various tasks such as named entity recognition, syntactic chunking or syntactic dependency parsing. It was also successfully applied for transfer learning of multilingual structure.

The most commonly used clustering method for semi-supervised learning is known as Brown clustering. While still being one of the most efficient word representation method, Brown clustering has two limitations we want to address in this work. First, since it is a hard clustering method, homonymy is ignored. Second, it does not take into account syntactic relations between words, which seems crucial to induce semantic classes. Our goal is thus to propose a method for semantic class induction which takes into account both syntax and homonymy, and then to study their effects on semantic class learning.

We start by introducing a new unsupervised method for semantic classes induction. This is achieved by defining a generative model of sentences with latent variables, which aims at capturing semantic roles of words. We require our method to be scalable, in order to learn models on large datasets containing tens of millions of sentences. More precisely, we make the following contributions:

- We introduce a generative model of sentences, based on dependency trees, which can be seen as a generalization of Brown clustering,
- We describe a fast approximate inference algorithm, based on message passing and online EM for scaling to large datasets. It allowed us to learn models with 512 latent states on a dataset with hundreds of millions of tokens in less than two days on a single core,
- We learn models on two datasets, Wikipedia articles about musicians and the NYT corpus, and evaluate them on two semi-supervised tasks, namely supersense tagging and named entity recognition.

6.20. Domain Adaptation for Sequence Labeling using Hidden Markov Models

Participants: Edouard Grave [correspondent], Guillaume Obozinski, Francis Bach.

Most natural language processing systems based on machine learning are not robust to domain shift. For example, a state-of-the-art syntactic dependency parser trained on Wall Street Journal sentences has an absolute drop in performance of more than ten points when tested on textual data from the Web. An efficient solution to make these methods more robust to domain shift is to first learn a word representation using large amounts of unlabeled data from both domains, and then use this representation as features in a supervised learning algorithm. In this paper, we propose to use hidden Markov models to learn word representations for part-of-speech tagging. In particular, we study the influence of using data from the source, the target or both domains to learn the representation and the different ways to represent words using an HMM.

Nowadays, most natural language processing systems are based on supervised machine learning. Despite the great successes obtained by those techniques, they unfortunately still suffer from important limitations. One of them is their sensitivity to domain shift: for example, a state-of-the-art part-of-speech tagger trained on the Wall Street Journal section of the Penn treebank achieves an accuracy of 97% when tested on sentences from the Wall Street Journal, but only 90% when tested on textual data from the Web. This drop in performance can also be observed for other tasks such as syntactic parsing or named entity recognition.

One of the explanations for this drop in performance is the big lexical difference that exists accross domains. This results in a lot of out-of-vocabulary words (OOV) in the test data, *i.e.*, words of the test data that were not observed in the training set. For example, more than 25% of the tokens of the test data from the Web corpus are unobserved in the training data from the WSJ. By comparison, only 11.5% of the tokens of the test data from the WSJ are unobserved in the training data from the WSJ. Part-of-speech taggers make most of their errors on those out-of-vocabulary words.

Labeling enough data to obtain a high accuracy for each new domain is not a viable solution. Indeed, it is expensive to label data for natural language processing, because it requires expert knowledge in linguistics. Thus, there is an important need for transfer learning, and more precisely for domain adaptation, in computational linguistics. A common solution consists in using large quantities of unlabeled data, from both source and target domains, in order to learn a good word representation. This representation is then used as features to train a supervised classifier that is more robust to domain shift. Depending on how much data from the source and the target domains are used, this method can be viewed as performing semi-supervised learning or domain adaptation. The goal is to reduce the impact of out-of-vocabulary words on performance. This scheme was first proposed to reduce data sparsity for named entity recognition, before being applied to domain adaptation for part-of-speech tagging or syntactic parsing.

Hidden Markov models have already been considered in previous work to learn word representations for domain adaptation or semi-supervised learning. Our contributions in [25] are mostly experimental: we compare different word representations that can be obtained from an HMM and study the effect of training the unsupervised HMM on source, target or both domains. While previous work mostly use Viterbi decoding to obtain word representations from an HMM, we empirically show that posterior distributions over latent classes give better results.

6.21. Simple Greedy Matching for Aligning Large Knowledge Bases

Participant: Simon Lacoste-Julien [correspondent].

Collaboration with: Konstantina Palla, Alex Davies, Zoubin Ghahramani (Machine Learning Group, Department of Engineering, University of Cambridge), Gjergji Kasneci (Max Planck Institut für Informatik), Thore Graepel (Microsoft Research Cambridge)

The Internet has enabled the creation of a growing number of large-scale knowledge bases in a variety of domains containing complementary information. Tools for automatically aligning these knowledge bases would make it possible to unify many sources of structured knowledge and answer complex queries. However, the efficient alignment of large-scale knowledge bases still poses a considerable challenge. Here, we present Simple Greedy Matching (SiGMa), a simple algorithm for aligning knowledge bases with millions of entities and facts. SiGMa is an iterative propagation algorithm that leverages both the structural information from the relationship graph and flexible similarity measures between entity properties in a greedy local search, which makes it scalable. Despite its greedy nature, our experiments in [17] indicate that SiGMa can efficiently match some of the world's largest knowledge bases with high accuracy. We provide additional experiments on benchmark datasets which demonstrate that SiGMa can outperform state-of-the-art approaches both in accuracy and efficiency.

SISYPHE Project-Team

6. New Results

6.1. Neuroscience & Neuroendocrinology: Regulation of the Gonadotrope axis

6.1.1. A numerical method for transport equations with discontinuous flux functions: application to mathematical modeling of cell dynamics

Participants: Benjamin Aymard, Frédérique Clément, Frédéric Coquel, Marie Postel.

We have proposed a numerical method to handle discontinuous fluxes arising in transport-like equations [35]. More precisely, we have studied hyperbolic PDEs with flux transmission conditions at interfaces between subdomains where coefficients are discontinuous. A dedicated finite volume scheme with a limited high order enhancement has been adapted to treat the discontinuities arising at interfaces. The validation of the method has been done on one- and two-dimensional toy problems for which exact solutions are available, allowing us to do a thorough convergence study. We have then applied the method to a biological model focusing on complex cell dynamics [40] that initially motivated this study and illustrates the full potentialities of the scheme.

6.1.2. Adaptive mesh refinement strategy for a nonconservative transport problem

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel.

In the framework of transport equations it is usual to need long time simulations, and therefore large physical domains to cover a phenomenon. On the other hand it can happen that only a small time varying portion of the domain is interesting. This motivates the use of adaptivity for the spatial discretization. Biological models involving cell development are often nonconservative to account for cell division. In that case the threshold controlling the spatial adaptivity may have to be time-dependent in order to keep up with the progression of the solution. We have tackled the difficulties arising when applying a multiresolution method to a transport equation with discontinuous fluxes modeling localized mitosis [76]. The analysis of the numerical method is performed on a simplified model and numerical scheme. An original threshold strategy is proposed and validated thanks to extensive numerical tests. It is then applied to a biological model in both cases of distributed and localized mitosis.

6.1.3. Coupled Somatic Cell Kinetics and Germ Cell Growth: Multiscale Model-Based Insight on Ovarian Follicular Development

Participants: Frédérique Clément, Philippe Michel, Danielle Monniaux, Thomas Stiehl.

We have designed a stochastic individual-based model describing the first stages of follicular development, where the cell population is structured with respect to age (progression within the cell cycle) and space (radial distance from the oocyte) [39]. The model accounts for the molecular dialogue existing between the oocyte and granulosa cells. Three dynamically interacting scales are considered in the model: (i) a microscopic, local scale corresponding to an individual cell embedded in its immediate environment, (ii) a mesoscopic, semi-local scale corresponding to anatomical or functional areas of follicles and (iii) a macroscopic, global scale corresponding to the morphology of the follicle. Numerical simulations were performed to reproduce the 3D morphogenesis of follicles and follow simultaneously the detailed spatial distribution of individual granulosa cells, their organization as concentric layers or functional cell clones and the increase in the follicle size. Detailed quantitative simulation results have been provided in the ovine species, in which well characterized genetic mutations lead to a variety of phenotypic follicle morphogenesis. The model can help to explain pathological situations of imbalance between oocyte growth and follicular cell proliferation.

6.1.4. Innovative computational and theoretical tools for slow-fast dynamics

Participants: Mathieu Desroches, Maciej Krupa.

Mixed-Mode Bursting Oscillations: Dynamics created by a slow passage through spike-adding canard explosion in a square-wave burster [44]. This work concerns the phenomenon of Mixed-Mode Bursting Oscillations (MMBOs). These are solutions of fast-slow systems of ordinary differential equations that exhibit both small-amplitude oscillations (SAOs) and bursts consisting of one or multiple large-amplitude oscillations (LAOs). The name MMBO is given in analogy to Mixed-Mode Oscillations, which consist of alternating SAOs and LAOs, without the LAOs being organized into burst events. In this article, we show how MMBOs are created naturally in systems that have a spike-adding bifurcation or spike-adding mechanism, and in which the dynamics of one (or more) of the slow variables causes the system to pass slowly through that bifurcation. Canards are central to the dynamics of MMBOs, and their role in shaping the MMBOs is two-fold: saddle-type canards are involved in the spike-adding mechanism of the underlying burster and permit one to understand the number of LAOs in each burst event, and folded-node canards arise due to the slow passage effect and control the number of SAOs. The analysis is carried out for a prototypical fourth-order system of this type, which consists of the third-order Hindmarsh-Rose system, known to have the spike-adding mechanism, and in which one of the key bifurcation parameters also varies slowly. We also include a discussion of the MMBO phenomenon for the Morris-Lecar-Terman system. Finally, we discuss the role of the MMBOs to a biological modeling of secreting neurons.

Canards in piecewise-linear systems: explosions and super-explosions [43]. We show that a planar slow-fast piecewise-linear (PWL) system with three zones admits limit cycles that share a lot of similarity with van der Pol canards, in particular an explosive growth. Using phase-space compactification, we show that these quasi-canard cycles are strongly related to a bifurcation at infinity. Furthermore, we investigate a limiting case in which we show the existence of a continuum of canard homoclinic connections that coexist for a single-parameter value and with amplitude ranging from an order of ε to an order of 1, a phenomenon truly associated with the non-smooth character of this system and which we call super-explosion.

Some results have been obtained concerning numerical continuation techniques for planar slow-fast systems [42] and short-term synaptic plasticity in the deterministic Tsodyks-Markram model that leads to unpredictable network dynamics [41].

6.2. Quantum engineering: controlled quantum systems

Participants: Joachim Cohen, Loïc Herviou, Mazyar Mirrahimi, Pierre Rouchon, Pierre Six.

6.2.1. Schrödinger cat states and hardware efficient quantum error correction

We introduce a new gate that transfers an arbitrary state of a qubit into a superposition of two quasi-orthogonal coherent states of a cavity mode, with opposite phases [52]. Such a highly non-classical state is often called a Schrödinger cat state. This qcMAP gate is based on conditional qubit and cavity operations exploiting the energy level dispersive shifts, in the regime where they are much stronger than the cavity and qubit linewidths. The generation of multi-component superpositions of quasi-orthogonal coherent states, non-local entangled states of two resonators and multi-qubit GHz states can be efficiently achieved by this gate.

In a second contribution [53], we propose to use an encoding of a quantum bit of information in a four-component Schrödinger cat state to ensure its protection against the photon loss, being the major source of decoherence for such a quantum harmonic oscillator. This protection is ensured by an efficient quantum error correction scheme employing the nonlinearity provided by a single physical qubit coupled to the cavity. We describe in detail how to implement these operations in a circuit quantum electrodynamics system. This directly addresses the task of building a hardware-efficient quantum memory and can lead to important shortcuts in quantum computing architectures.

As an important step towards the realization of such a protected quantum memory, in a collaboration with the team of Robert J. Schoelkopf at Yale university, we have successfully realized the encoding protocol of [52] using a 3D transom qubit coupled to a waveguide cavity resonator with a highly ideal off-resonant coupling [60]. This dispersive interaction is much greater than decoherence rates and higher-order nonlinearities to allow simultaneous manipulation of hundreds of photons. We created cat states as large as 111 photons and created superpositions of up to four coherent states. This control creates a powerful interface between discrete and continuous variable quantum computation and could enable applications in metrology and quantum information processing. This important achievement was published in Science and was also highlighted in Science Perspectives [103].

6.2.2. Quantum reservoir (dissipation) engineering

We have studied the application of dissipation engineering techniques to perform a high-performance and fast qubit reset [64]. Qubit reset is crucial at the start of and during quantum information algorithms. In a collaboration with the team of Michel H. Devoret at Yale university, our protocol, nicknamed DDROP (Double Drive Reset of Population) was experimentally tested on a superconducting transmon qubit and achieves a ground state preparation of at least 99.5% in times less than $3\mu s$; faster and higher fidelity are predicted upon parameter optimization [46].

Next, we proposed a dissipation engineering scheme that prepares and protects a maximally entangled state of a pair of superconducting qubits [54]. This is done by off-resonantly coupling the two qubits to a low-Q cavity mode playing the role of a dissipative reservoir. We engineer this coupling by applying six continuous-wave microwave drives with appropriate frequencies. The two qubits need not be identical. We show that our approach does not require any fine-tuning of the parameters and requires only that certain ratios between them be large. This protocol was experimentally realized in a collaboration with the team of M. H. Devoret at Yale university [57]. Unlike conventional, measurement-based schemes, this autonomous approach uses engineered dissipation to counteract decoherence, obviating the need for a complicated external feedback loop to correct errors. Instead, the feedback loop is built into the Hamiltonian such that the steady state of the system in the presence of drives and dissipation is a Bell state, an essential building block for quantum information processing. Such autonomous schemes, which are broadly applicable to a variety of physical systems, will be an essential tool for the implementation of quantum error correction. This important result appeared in Nature back-to-back to another paper by the group of D.J. Wineland (2012 Nobel prize winner) at NIST implementing similar ideas on another physical system consisting of trapped ion qubits [105].

6.2.3. Quantum measurement and measurement-based feedback

Measuring a quantum system can randomly perturb its state. The strength and nature of this back-action depend on the quantity that is measured. In a partial measurement performed by an ideal apparatus, quantum physics predicts that the system remains in a pure state whose evolution can be tracked perfectly from the measurement record. This property was proved in a collaboration with the group of Michel H. Devoret (Yale university) using a superconducting qubit dispersively coupled to a cavity traversed by a microwave signal [47]. The back-action on the qubit state of a single measurement of both signal quadratures was observed and shown to produce a stochastic operation whose action is determined by the measurement result. This accurate monitoring of a qubit state is an essential prerequisite for measurement-based feedback control of quantum systems. Indeed, in another experiment performed by our collaborators at ENS (team of Benjamin Huard and François Mallet), we demonstrated stabilization of an arbitrary trajectory of a superconducting qubit by such a measurement-based feedback [37]. The protocol benefits from the long coherence time ($T_2 > 10\mu$s) of the 3D transmon qubit, the high efficiency (82%) of the phase preserving Josephson amplifier, and fast electronics ensuring less than 500 ns delay. At discrete time intervals, the state of the qubit is measured and corrected in case an error is detected. For Rabi oscillations, where the discrete measurements occur when the qubit is supposed to be in the measurement pointer states, we demonstrate an average fidelity of 85% to the targeted trajectory. Incidentally, we demonstrate a fast reset protocol allowing to cool a 3D transmon qubit down to 0.6% in the excited state.

6.3. Classical engineering: Monitoring and control of complex systems

6.3.1. Modeling, signal analysis and control with medical applications

Participants: Alexandre Guerrini, Lisa Guigue, Claire Médigue, Michel Sorine, Serge Steer.

Reduced order cardiac modeling and applications. See Section 4.3 for complements. We consider two topics:
- Personalized medecine: a first validation on clinical data of our model of controlled contraction of cardiac muscle has been obtained [55].
- Heart Failure with preserved Ejection Fraction (HFpEF): this work is done in collaboration with Bijan Gahleh (INSERM U955). Our objective is to define markers of HFpEF identifiable from noninvasive measurements. After having assembled a high precision ECG acquisition and post-processing system, we have measured multi-lead ECG on pigs treated to induce HFpEF, cf. B. Gahleh et al [109]. The analysis of the diastolic electric interval (e.g. P-wave, PR interval etc.) is ongoing.

Semiclassical analysis of cardiovascular signals. A summary of the theory is now published [51].

CGAO-REA: Computerized Glucose Control in Critically Ill Patients. The version CGAO_v1 of our controller (see Sections 4.3), has been used in a large multi-center study, CGAO-REA (35 active ICUs, more than 3500 included patients). Mortality has not been changed [70], [48] but the protocol is now formalized and tunable. CGAO-REA has proved that our controller is robust in the real life context and comparable to human control with its present tuning. Improving the tuning (in particular the glycemic target) seems possible.

6.3.2. Diagnosis of inhomogeneous insulation degradation in electric cables by distributed shunt conductance estimation

Participant: Qinghua Zhang.

For the diagnosis of inhomogeneous insulation degradation in electric cables, the estimation of distributed shunt conductance is studied in this work. Gradual growth of the shunt conductance is a consequence of degradation of the dielectric properties of the insulator. The proposed estimation method is based on voltage and current measurements at a single end of the cable. After the linearization of the bilinear term of the telegrapher's equations through a perturbation approach, the Kalman filter is applied to transform the problem of dynamic system parameter estimation to a simple linear regression problem. Numerical simulations are made to demonstrate the feasibility of the proposed method. In particular, it is shown that the weak sensitivity of the available measurements to the shunt conductance can be compensated by long time data samples. See [61] for more details.

6.3.3. Feasibility of reflectometry techniques for non destructive evaluation of external post-tensioned cables

Participants: Michel Sorine, Qinghua Zhang.

Nowadays a considerable number of bridges is reaching an age when renovating operations become necessary. For some bridges, external post-tension is realized with cables protected in ducts, with the residual internal space imperfectly filled with a fluid cement grout. Detecting the problems of injection in the ducts is visually impossible from the outside. In collaboration with IFSTTAR (Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux) through the I4S team common to Inria and IFSTTAR, the feasibility of reflectometry techniques for cable health monitoring is investigated via numerical simulations and laboratory experiments. The main idea consists in adding electrically conductive tapes along a duct so that the duct and the added tapes can be treated as an electrical transmission line. It is then possible to apply advanced reflectometry methods developed by the SISYPHE project-team, initially for true electric cables.

6.3.4. Nonlinear system identification

Participants: Boyi Ni, Michel Sorine, Qinghua Zhang.

In the framework of the joint Franco-Chinese ANR-NSFC EBONSI project (see Section 8.1.1), the topics studied this year on nonlinear system identification are mainly on the detection of asymmetric control valve stiction from oscillatory data based on a method for extended Hammerstein system identification, and on the identification of Wiener systems.

The study on control valve stiction is motivated by the detection of control valves with asymmetric stiction resulting in oscillations in feedback control loops. The joint characterization of the control valve and the controlled process is formulated as the identification of a class of extended Hammerstein systems. The input nonlinearity is described by a point-slope-based hysteretic model with two possibly asymmetric ascent and descent paths. An iterative identification method is proposed, based on the idea of separating the ascent and descent paths subject to the oscillatory input and output. The structure of the formulated extended Hammerstein system is shown to be identifiable, and the oscillatory signals in feedback control loops are proved to be informative by exploiting the cyclo-stationarity of these oscillatory signals. Numerical, experimental and industrial examples confirm the effectiveness of the proposed identification method.

Wiener system identification has been investigated this year by focusing on the estimation of the finite impulse response (FIR) of the linear subsystem. Under the assumption of Gaussian input distribution, this work mainly aims at addressing a deficiency of the well-known correlation-based method for Wiener system identification: it fails when the nonlinearity of the Wiener system is an even function. This method is, in the considered Gaussian input case, equivalent to the best linear approximation (BLA), which exhibits the same deficiency. A new method is developed this year, based on a weighted principal component analysis (wPCA). Its consistency is proved for Wiener systems with either even or non even nonlinearities. Its computational cost is almost the same as that of a standard PCA. Numerical simulations are made to compare the new wPCA-based method to the correlation-based method for different Wiener systems with nonlinearities more or less close to an even function.

6.3.5. Model-based fault diagnosis for descriptor systems

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

This work is about fault diagnosis for linear time varying descriptor systems, the discrete time counterpart of dynamic systems described by differential-algebraic equations. The Kalman filter for descriptor systems is first revisited by completing existing results about its properties that are essential for the purpose of fault diagnosis. Based on the analysis of the effects of the considered actuator and sensor faults on the innovation of the Kalman filter, it is shown that the considered fault diagnosis problem in linear time varying descriptor systems can be transformed to a classical linear regression problem formulated by appropriately filtering the input-output data. Following this result, algorithms for fault diagnosis through maximum likelihood estimation are then developed.

In the framework of the ITEA2 MODRIO project (see Section 8.2.1), this work is in preparation for studying hybrid system monitoring, aiming at extending existing results from state-space systems to descriptor systems in the modes of a hybrid system.

6.3.6. Analysis of the Behavior of Networks of Dynamical Systems

Participant: Pierre-Alexandre Bliman.

We have established convergence results for some continuous-time dynamics which are analogs to ant colony optimization algorithms that solve shortest path problems. Global asymptotic stability has been shown, and the speed of convergence has been calculated explicitly and shown to be proportional to the difference between the reciprocals of the second shortest and the shortest paths. Such precise results are missing in the context of ant colony optimization algorithms (which are discrete-time dynamical systems). The systems studied are special instances of networks of dynamical systems which represent the evolution of some state variable on each path, coupled in a competitive way through global macroscopic quantity. Such models are related to simple forms of models studied in mathematical epidemiology, which will be the subject of further work. This work is done in cooperation with Amit Bhaya from COPPE, Universidade Federal de Rio de Janeiro. Papers have been submitted [77].

SMIS Project-Team

6. New Results

6.1. Minimum Exposure

Participants: Nicolas Anciaux, Marouane Fazouane, Benjamin Nguyen [correspondent], Michalis Vazirgiannis.

When users request a service, the service provider usually asks for personal documents to tailor its service to the specific situation of the applicant. For example, the rate and duration of consumer's loans are usually adapted depending on the risk based on the income, assets or past lines of credits of the borrower. In practice, an excessive amount of personal data is collected and stored. Indeed, a paradox is at the root of this problem: service providers require users to expose data in order to determine whether that data is needed or not to achieve the purpose of the service. We explore a reverse approach, where service providers would publicly describe the data they require to complete their task, and where software (placed, depending on the context, on the client, on the server, or in a trusted hardware component) would use those descriptions to determine a minimum subset of information to expose.

Following our 2012 seminal works on the general Minimum Exposure framework, we have pursued its general study in 2013 [15], [29]. We have also developed a prototype system, using a low powered and highly secure smartcard [21], which is used to support hidden decision rules.

6.2. Flash-Based Data Management

Participants: Nicolas Anciaux, Matias Bjørling, Philippe Bonnet, Luc Bouganim [correspondent], Niv Dayan, Philippe Pucheral.

Mass-storage secure portable tokens are emerging and provide a real breakthrough in the management of sensitive data. They can embed personal data and/or metadata referencing documents stored encrypted in the Cloud and can manage them under holder's control. Mass on-board storage requires efficient embedded database techniques. These techniques are however very challenging to design due to a combination of conflicting NAND Flash constraints and scarce RAM constraint, disqualifying known state of the art solutions. To tackle this challenge, we proposed a log-only based storage organization and an appropriate indexing scheme, which (1) produce only sequential writes compatible with the Flash constraints and (2) consume a tiny amount of RAM, independent of the database size [13].

Solid State Drives (SSDs) are a moving target for system designers: they are black boxes, their internals are undocumented, and their performance characteristics vary across models. There is no appropriate analytical model and experimenting with commercial SSDs is cumbersome, as it requires a careful experimental methodology to ensure repeatability. Worse, performance results obtained on a given SSD cannot be generalized. Overall, it is impossible to explore how a given algorithm, say a hash join or LSM-tree insertions, leverages the intrinsic parallelism of a modern SSD, or how a slight change in the internals of an SSD would impact its overall performance. In 2013, we worked on a new SSD simulation framework, named EagleTree, which addresses these problems, and enables a principled study of SSD-Based algorithms. We published a demonstration on EagleTree at VLDB'13 [20]. The demonstration scenario illustrates the design space for algorithms based on an SSD-based IO stack, and shows how researchers and practitioners can use EagleTree to perform tractable explorations of this complex design space.

6.3. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Philippe Pucheral, Cuong Quoc To.

Current applications, from complex sensor systems (e.g. quantified self) to online e-markets acquire vast quantities of personal information which usually ends-up on central servers. Decentralized architectures, devised to help individuals keep full control of their data, hinder global treatments and queries, impeding the development of services of great interest. In this study, we promote the idea of pushing the security to the edges of applications, through the use of secure hardware devices controlling the data at the place of their acquisition. To solve this problem, we propose secure distributed querying protocols based on the use of a tangible physical element of trust, reestablishing the capacity to perform global computations without revealing any sensitive information to central servers. This leads to execute global treatments on an asymmetric architecture, composed of a powerful, available and untrusted computing infrastructure (server or cloud), and a large set of low powered, highly disconnected trusted devices. Given our large scale data centric applications (e.g. nationwide surveys), we discard solutions based on secure multi-party computation, which do not scale. We have primarily studied 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 [12]. This work is an extension of [26]. A vulgarization paper on the scientific and societal challenges related to PPDP techniques has been published in a newspaper [24]. We are now trying to support general SQL queries in this same execution context. We concentrate first 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. Cost models and experiments demonstrate that this approach can scale to nationwide infrastructures [23], [42]. This work is part of Cuong Quoc To's Ph.D. thesis started in sept. 2012.

6.4. Trusted Cells

Participants: Nicolas Anciaux, Philippe Bonnet, Luc Bouganim, Benjamin Nguyen, Pilippe Pucheral [correspondent], Iulian Sandu Popa.

With the convergence of mobile communications, sensors and online social networks technologies, we are witnessing an exponential increase in the creation and consumption of personal data. Such data is volunteered by users, automatically captured by sensors or inferred from existing data. Today, there is a wide consensus that individuals should have increased control on how their personal data is collected, managed and shared. Yet there is no appropriate technical solution to implement such personal data services: centralized solutions sacrifice security for innovative applications, while decentralized solutions sacrifice innovative applications for security. In this work, we argue that the advent of secure hardware in all personal IT devices, at the edges of the Internet, could trigger a sea change. We propose the vision of trusted cells: personal data servers running on secure smart phones, set-top boxes, secure portable tokens or smart cards to form a global, decentralized data platform that provides security yet enables innovative applications. We motivate our approach, describe the trusted cells architecture and define a range of challenges for future research in a paper published at CIDR'13 (Int. Conf on Innovative Data Systems Research). This work was based on a thorough analysis of existing and potential threats on personal data, which led to a tutorial on data privacy [18], [30].

In parallel, we revisited the Trusted Cells vision to the context of Least Developed Countries (LDCs). The main barrier to the development of IT services in these regions is not only the lack of communication facilities, but also the lack of consistent information systems, security procedures, economic and legal support. We propose, Folk-enabled Information System (Folk-IS), a new paradigm based on a fully decentralized and participatory approach, where each individual implements a small subset of a complete information system without the need for infrastructure. As trusted cells, Folk-IS builds upon the emergence of highly secure, portable, low-cost storage and computing devices, called hereafter Smart Tokens. Here, however, the focus is on the low cost of ownership, deployment and maintenance, and on the absence of a networked infrastructure. With Folk-IS, and thanks to 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 [17].

WILLOW Project-Team

6. New Results

6.1. 3D object and scene modeling, analysis, and retrieval

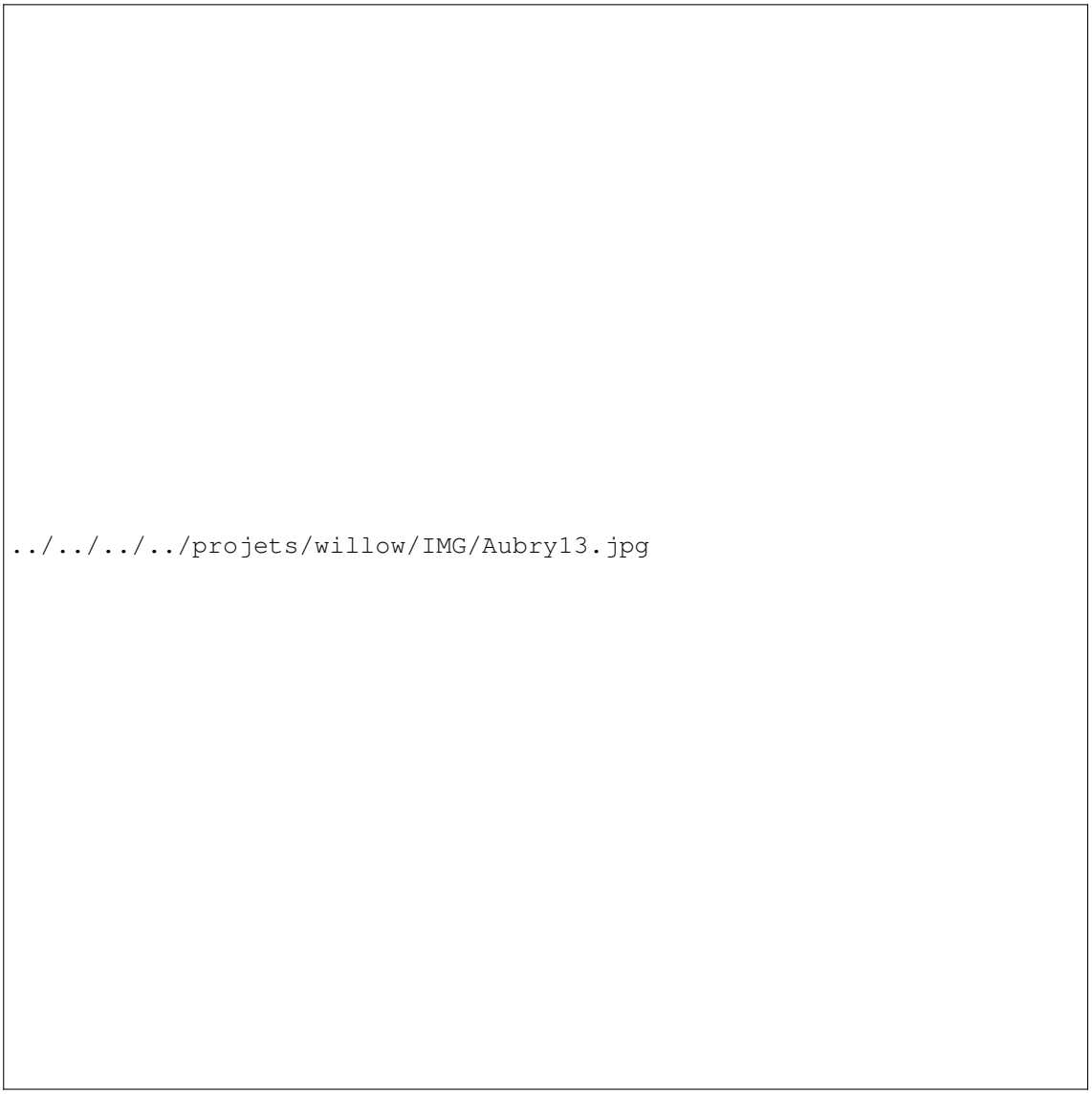


Figure 1. Our system automatically aligns and recovers the viewpoint of paintings, drawings, and historical photographs to a 3D model of an architectural site.

6.1.1. Painting-to-3D Model Alignment Via Discriminative Visual Elements

Participants: Mathieu Aubry, Bryan Russell [Intel Labs], Josef Sivic.

In this work we describe a technique that can reliably align arbitrary 2D depictions of an architectural site, including drawings, paintings and historical photographs, with a 3D model of the site. This is a tremendously difficult task as the appearance and scene structure in the 2D depictions can be very different from the appearance and geometry of the 3D model, e.g., due to the specific rendering style, drawing error, age, lighting or change of seasons. In addition, we face a hard search problem: the number of possible alignments of the painting to a large 3D model, such as a partial reconstruction of a city, is huge. To address these issues, we develop a new compact representation of complex 3D scenes. The 3D model of the scene is represented by a small set of discriminative visual elements that are automatically learnt from rendered views. Similar to object detection, the set of visual elements, as well as the weights of individual features for each element, are learnt in a discriminative fashion. We show that the learnt visual elements are reliably matched in 2D depictions of the scene despite large variations in rendering style (e.g. watercolor, sketch, historical photograph) and structural changes (e.g. missing scene parts, large occluders) of the scene. We demonstrate an application of the proposed approach to automatic re-photography to find an approximate viewpoint of historical paintings and photographs with respect to a 3D model of the site. The proposed alignment procedure is validated via a human user study on a new database of paintings and sketches spanning several sites. The results demonstrate that our algorithm produces significantly better alignments than several baseline methods. This work has been accepted for publication to the ACM Transactions on Graphics (ACM ToG). The problem addressed in this work is illustrated in Figure 1 and example results are shown in figure 2 . The pre-print is available online at [10].

6.1.2. Learning and Calibrating Per-Location Classifiers for Visual Place Recognition

Participants: Petr Gronat, Josef Sivic, Guillaume Obozinski [ENPC / Inria SIERRA], Tomáš Pajdla [CTU in Prague].

The aim of this work is to localize a query photograph by finding other images depicting the same place in a large geotagged image database. This is a challenging task due to changes in viewpoint, imaging conditions and the large size of the image database. The contribution of this work is two-fold. First, we cast the place recognition problem as a classification task and use the available geotags to train a classifier for each location in the database in a similar manner to per-exemplar SVMs in object recognition. Second, as only few positive training examples are available for each location, we propose a new approach to calibrate all the per-location SVM classifiers using *only* the negative examples. The calibration we propose relies on a significance measure essentially equivalent to the p-values classically used in statistical hypothesis testing. Experiments are performed on a database of 25,000 geotagged street view images of Pittsburgh and demonstrate improved place recognition accuracy of the proposed approach over the previous work. This work has been published at CVPR 2013 [6].

6.1.3. Visual Place Recognition with Repetitive Structures

Participants: Akihiko Torii [Tokyo Institute of Technology], Josef Sivic, Tomáš Pajdla [CTU in Prague], Masatoshi Okutomi [Tokyo Institute of Technology].

Repeated structures such as building facades, fences or road markings often represent a significant challenge for place recognition. Repeated structures are notoriously hard for establishing correspondences using multi-view geometry. Even more importantly, they violate the feature independence assumed in the bag-of-visual-words representation which often leads to over-counting evidence and significant degradation of retrieval performance. In this work we show that repeated structures are not a nuisance but, when appropriately represented, they form an important distinguishing feature for many places. We describe a representation of repeated structures suitable for scalable retrieval. It is based on robust detection of repeated image structures and a simple modification of weights in the bag-of-visual-word model. Place recognition results are shown on datasets of street-level imagery from Pittsburgh and San Francisco demonstrating significant gains in recognition performance compared to the standard bag-of-visual-words baseline and more recently proposed burstiness weighting. This work has been published at CVPR 2013 [9].

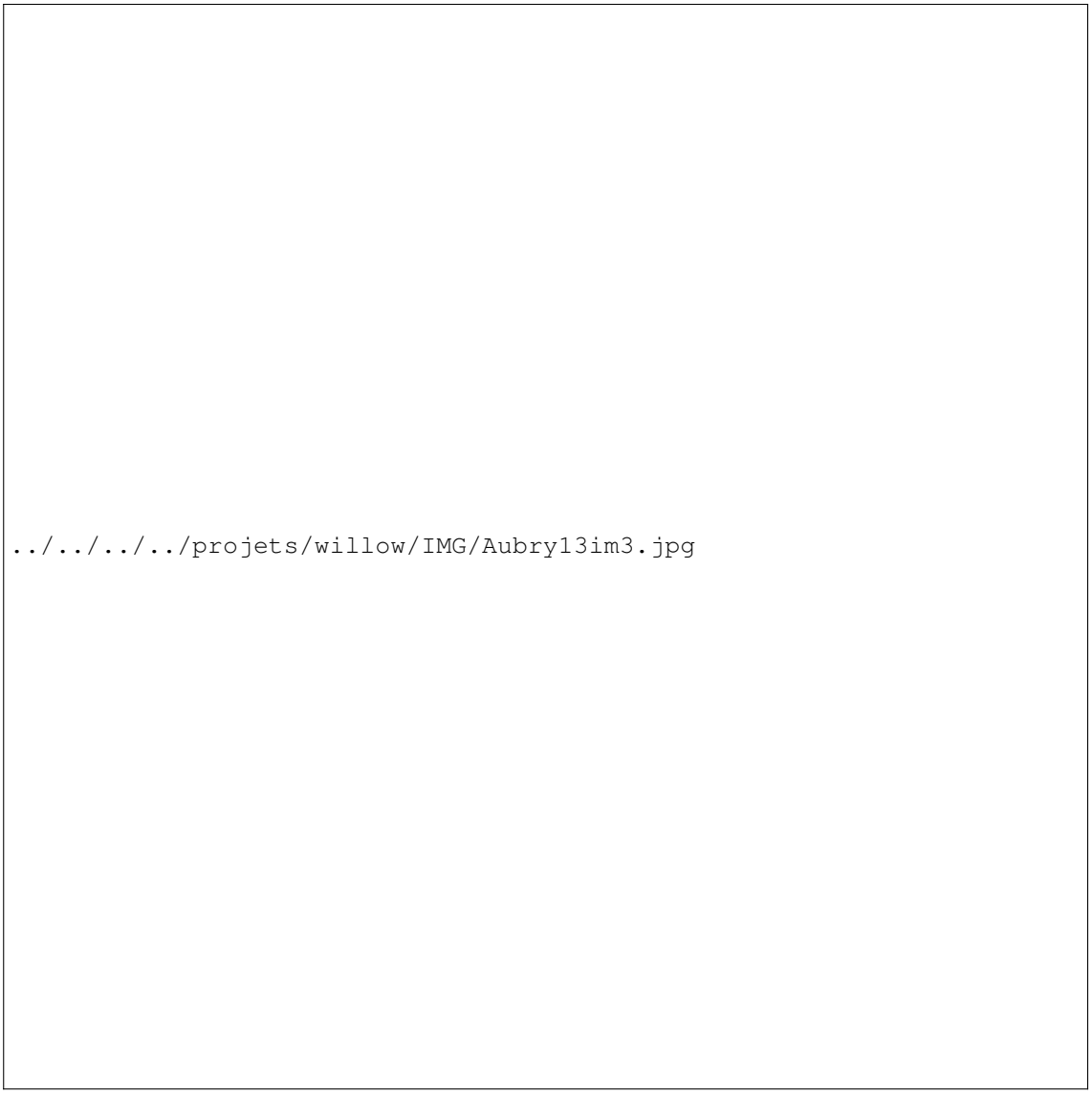


Figure 2. Example alignments of non-photographic depictions to 3D models. Notice that we are able to align depictions rendered in different styles and having a variety of viewpoints with respect to the 3D models.

6.1.4. Trinocular Geometry Revisited

Participants: Jean Ponce, Martial Hebert [CMU].

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. This work has been submitted to CVPR 2014.

6.2. Category-level object and scene recognition

6.2.1. Learning Graphs to Match

Participants: Minsu Cho, Karteek Alahari, Jean Ponce.

Many tasks in computer vision are formulated as graph matching problems. Despite the NP-hard nature of the problem, fast and accurate approximations have led to significant progress in a wide range of applications. Learning graph models from observed data, however, still remains a challenging issue. This work presents an effective scheme to parameterize a graph model, and learn its structural attributes for visual object matching. For this, we propose a graph representation with histogram-based attributes, and optimize them to increase the matching accuracy. Experimental evaluations on synthetic and real image datasets demonstrate the effectiveness of our approach, and show significant improvement in matching accuracy over graphs with pre-defined structures. The work is illustrated in Figure 3 . This work has been published ICCV 2013 [3].

6.2.2. Finding Matches in a Haystack: A Max-Pooling Strategy for Graph Matching in the Presence of Outliers

Participants: Minsu Cho, Olivier Duchenne [Intel], Jian Sun, Jean Ponce.

A major challenge in real-world matching problems is to tolerate the numerous outliers arising in typical visual tasks. Variations in object appearance, shape, and structure within the same object class make it hard to distinguish inliers from outliers due to clutters. In this project, we propose a novel approach to graph matching, which is not only resilient to deformations but also remarkably tolerant to outliers. By adopting a max-pooling strategy within the graph matching framework, the proposed algorithm evaluates each candidate match using its most promising neighbors, and gradually propagates the corresponding scores to update the neighbors. As final output, it assigns a reliable score to each match together with its supporting neighbors, thus providing contextual information for further verification. We demonstrate the robustness and utility of our method with synthetic and real image experiments. This work has been submitted to CVPR 2014.

6.2.3. Decomposing Bag of Words Histograms

Participants: Ankit Gandhi [IIIT India], Karteek Alahari, C.v. Jawahar [IIIT India].

We aim to decompose a global histogram representation of an image into histograms of its associated objects and regions. This task is formulated as an optimization problem, given a set of linear classifiers, which can effectively discriminate the object categories present in the image. Our decomposition bypasses harder problems associated with accurately localizing and segmenting objects. We evaluate our method on a wide variety of composite histograms, and also compare it with MRF-based solutions. In addition to merely measuring the accuracy of decomposition, we also show the utility of the estimated object and background histograms for the task of image classification on the PASCAL VOC 2007 dataset. This work has been published at ICCV 2013 [5].

6.2.4. Image Retrieval using Textual Cues

Participants: Anand Mishra [IIIT India], Karteek Alahari, C.v. Jawahar [IIIT India].

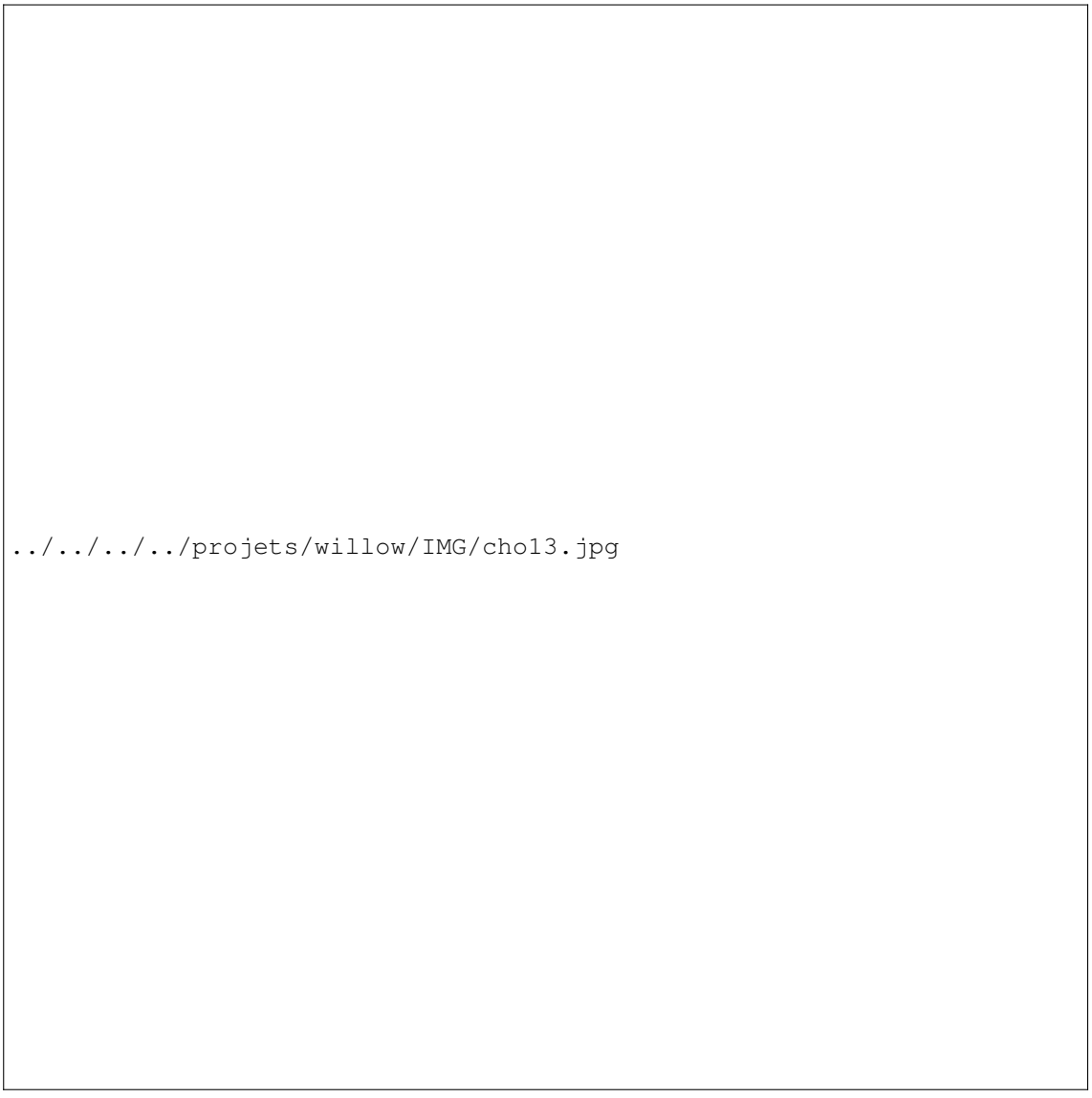


Figure 3. Graph learning for matching. Our approach learns a graph model from labeled data to provide the best match to instances of a target class. It shows significant improvement over previous approaches for matching. (Best viewed in color.)

We present an approach for the text-to-image retrieval problem based on textual content present in images. Given the recent developments in understanding text in images, an appealing approach to address this problem is to localize and recognize the text, and then query the database, as in a text retrieval problem. We show that such an approach, despite being based on state-of-the-art methods, is insufficient, and propose a method, where we do not rely on an exact localization and recognition pipeline. We take a query-driven search approach, where we find approximate locations of characters in the text query, and then impose spatial constraints to generate a ranked list of images in the database. The retrieval performance is evaluated on public scene text datasets as well as three large datasets, namely IIIT scene text retrieval, Sports-10K and TV series-1M, we introduce. This work has been published at ICCV 2013 [7].

6.2.5. *Learning Discriminative Part Detectors for Image Classification and Cosegmentation*

Participants: Jian Sun, Jean Ponce.

In this work, we address the problem of learning discriminative part detectors from image sets with category labels. We propose a novel latent SVM model regularized by group sparsity to learn these part detectors. Starting from a large set of initial parts, the group sparsity regularizer forces the model to jointly select and optimize a set of discriminative part detectors in a max-margin framework. We propose a stochastic version of a proximal algorithm to solve the corresponding optimization problem. We apply the proposed method to image classification and cosegmentation, and quantitative experiments with standard bench- marks show that it matches or improves upon the state of the art. This work has been published at CVPR 2013 [8].

6.2.6. *Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks*

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

Convolutional neural networks (CNN) have recently shown outstanding image classification performance in the large-scale visual recognition challenge (ILSVRC2012). The success of CNNs is attributed to their ability to learn rich mid-level image representations as opposed to hand-designed low-level features used in other image classification methods. Learning CNNs, however, amounts to estimating millions of parameters and requires a very large number of annotated image samples. This property currently prevents application of CNNs to problems with limited training data. In this work we show how image representations learned with CNNs on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks with limited amount of training data. We design a method to reuse layers trained on the ImageNet dataset to compute mid-level image representation for images in the PASCAL VOC dataset. We show that despite differences in image statistics and tasks in the two datasets, the transferred representation leads to significantly improved results for object and action classification, outperforming the current state of the art on Pascal VOC 2007 and 2012 datasets. We also show promising results for object and action localization. The pre-print of this work is available online [11]. Results are illustrated in Figure 4 .

6.2.7. *Seeing 3D chairs: exemplar part-based 2D-3D alignment using a large dataset of CAD models*

Participants: Mathieu Aubry, Bryan Russell [Intel labs], Alyosha Efros [UC Berkeley], Josef Sivic.

We present an approach for the text-to-image retrieval problem based on textual content present in images. Given the recent developments in understanding text in images, an appealing approach to address this problem is to localize and recognize the text, and then query the database, as in a text retrieval problem. We show that such an approach, despite being based on state-of-the-art methods, is insufficient, and propose a method, where we do not rely on an exact localization and recognition pipeline. We take a query-driven search approach, where we find approximate locations of characters in the text query, and then impose spatial constraints to generate a ranked list of images in the database. The retrieval performance is evaluated on public scene text datasets as well as three large datasets, namely IIIT scene text retrieval, Sports-10K and TV series-1M, we introduce. This work has been submitted to CVPR 2014.

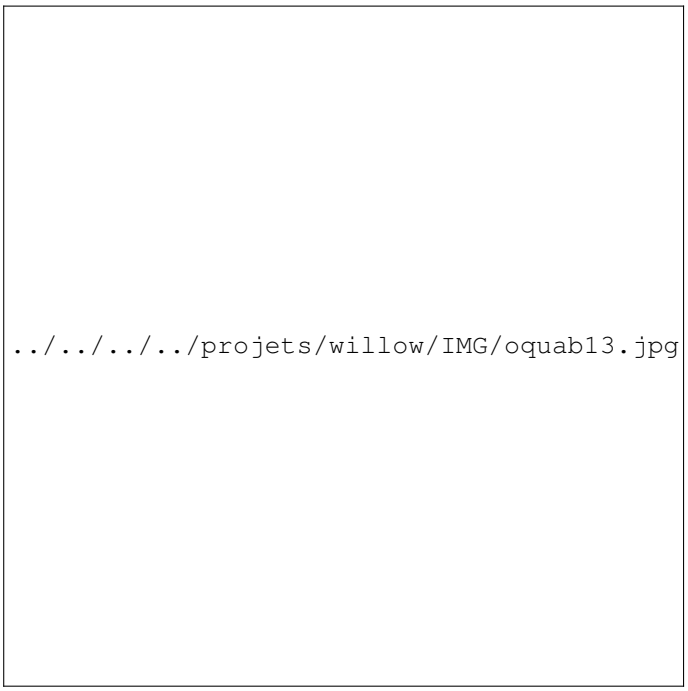


Figure 4. Recognition and localization results of our method for a Pascal VOC test image. Output maps are shown for six object categories with the highest responses.

6.3. Image restoration, manipulation and enhancement

6.3.1. Learning to Estimate and Remove Non-uniform Image Blur

Participants: Florent Couzinie-Devy, Jian Sun, Karteek Alahari, Jean Ponce.

This work addresses the problem of restoring images subjected to unknown and spatially varying blur caused by defocus or linear (say, horizontal) motion. The estimation of the global (non-uniform) image blur is cast as a multi-label energy minimization problem. The energy is the sum of unary terms corresponding to learned local blur estimators, and binary ones corresponding to blur smoothness. Its global minimum is found using Ishikawa's method by exploiting the natural order of discretized blur values for linear motions and defocus. Once the blur has been estimated, the image is restored using a robust (non-uniform) deblurring algorithm based on sparse regularization with global image statistics. The proposed algorithm outputs both a segmentation of the image into uniform-blur layers and an estimate of the corresponding sharp image. We present qualitative results on real images, and use synthetic data to quantitatively compare our approach to the publicly available implementation of Chakrabarti et al. 2010. This work has been published at CVPR 2013 [4] and example results are shown in figure 5 .

6.3.2. Efficient, Blind, Spatially-Variant Deblurring for Shaken Images

Participants: Oliver Whyte [Microsoft Redmond], Josef Sivic, Andrew Zisserman, Jean Ponce.

In this chapter we discuss modeling and removing spatially-variant blur from photographs. We describe a compact global parameterization of camera shake blur, based on the 3D rotation of the camera during the exposure. Our model uses three-parameter homographies to connect camera motion to image motion and, by assigning weights to a set of these homographies, can be seen as a generalization of the standard, spatially-invariant convolutional model of image blur. As such we show how existing algorithms, designed for spatially-invariant deblurring, can be "upgraded" in a straightforward manner to handle spatially-variant blur instead. We demonstrate this with algorithms working on real images, showing results for blind estimation of blur parameters from single images, followed by non-blind image restoration using these parameters. Finally, we introduce an efficient approximation to the global model, which significantly reduces the computational cost of modeling the spatially-variant blur. By approximating the blur as locally-uniform, we can take advantage of fast Fourier-domain convolution and deconvolution, reducing the time required for blind deblurring by an order of magnitude.

This work has been accepted for publication as a book chapter in the upcoming book "Motion Deblurring: Algorithms and Systems" to be published by Cambridge University Press in May 2014. [2] The demo implementing deblurring of images degraded by camera shake is available online at: http://www.di.ens.fr/willow/research/saturation/.

6.4. Human activity capture and classification

6.4.1. Layered Segmentation of People in Stereoscopic Movies

Participants: Karteek Alahari, Guillaume Seguin, Josef Sivic, Ivan Laptev.

In this work we seek to obtain a pixel-wise segmentation and pose estimation of multiple people in a stereoscopic video. This involves challenges such as dealing with unconstrained stereoscopic video, non-stationary cameras, and complex indoor and outdoor dynamic scenes. The contributions of our work are two-fold: First, we develop a segmentation model incorporating person detection, pose estimation, as well as colour, motion, and disparity cues. Our new model 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 2727 realistic stereo pairs and includes annotation of human poses, person bounding boxes, and pixel-wise segmentations for hundreds 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). This work has been published at ICCV 2013 [1].

[2] http://www.cambridge.org/fr/academic/subjects/engineering/image-processing-and-machine-vision/motion-deblurring-algorithms-and-systems

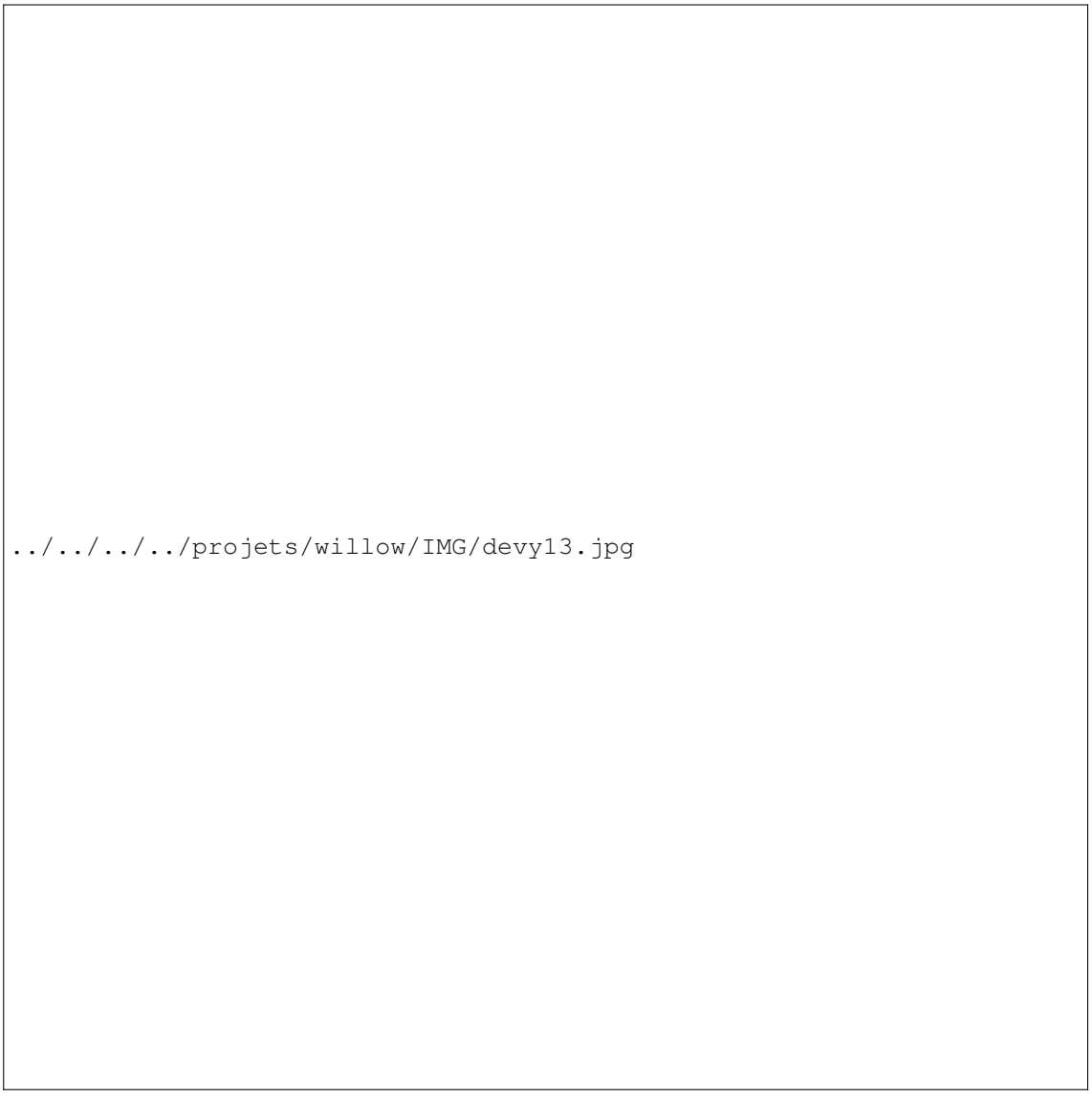


Figure 5. *Sample deblurring results on real images. From left to right: blurry image, deblurred image, close-up corresponding to the boxes shown in red. Note that our estimated deblurred image has more detail.*

6.4.2. Finding Actors and Actions in Movies

Participants: Piotr Bojanowski, Francis Bach [Inria Sierra], Ivan Laptev, Jean Ponce, Cordelia Schmid [Inria Lear], Josef Sivic.

We address the problem of learning a joint model of actors and actions in movies using weak supervision provided by scripts. Specifically, we extract actor/action pairs from the script and use them as constraints in a discriminative clustering framework. The corresponding optimization problem is formulated as a quadratic program under linear constraints. People in video are represented by automatically extracted and tracked faces together with corresponding motion features. First, we apply the proposed framework to the task of learning names of characters in the movie and demonstrate significant improvements over previous methods used for this task. Second, we explore the joint actor/action constraint and show its advantage for weakly supervised action learning. We validate our method in the challenging setting of localizing and recognizing characters and their actions in feature length movies Casablanca and American Beauty. This work has been published at ICCV 2013 [2] and example results are shown in figure 6 . The corresponding software has been also made publicly available (see the software section of this report).

6.4.3. Highly-Efficient Video Features for Action Recognition and Counting

Participants: Vadim Kantorov, Ivan Laptev.

Local video features provide state-of-the-art performance for action recognition. While the accuracy of action recognition has been steadily improved over the recent years, the low speed of feature extraction remains to be a major bottleneck preventing current methods from addressing large-scale applications. In this work we demonstrate that local video features can be computed very efficiently by exploiting motion information readily-available from standard video compression schemes. We show experimentally that the use of sparse motion vectors provided by the video compression improves the speed of existing optical-flow based methods by two orders of magnitude while resulting in limited drops of recognition performance. Building on this representation, we next address the problem of event counting in video and present a method providing accurate counts of human actions and enabling to process 100 years of video on a modest computer cluster. This work has been submitted to CVPR 2014.

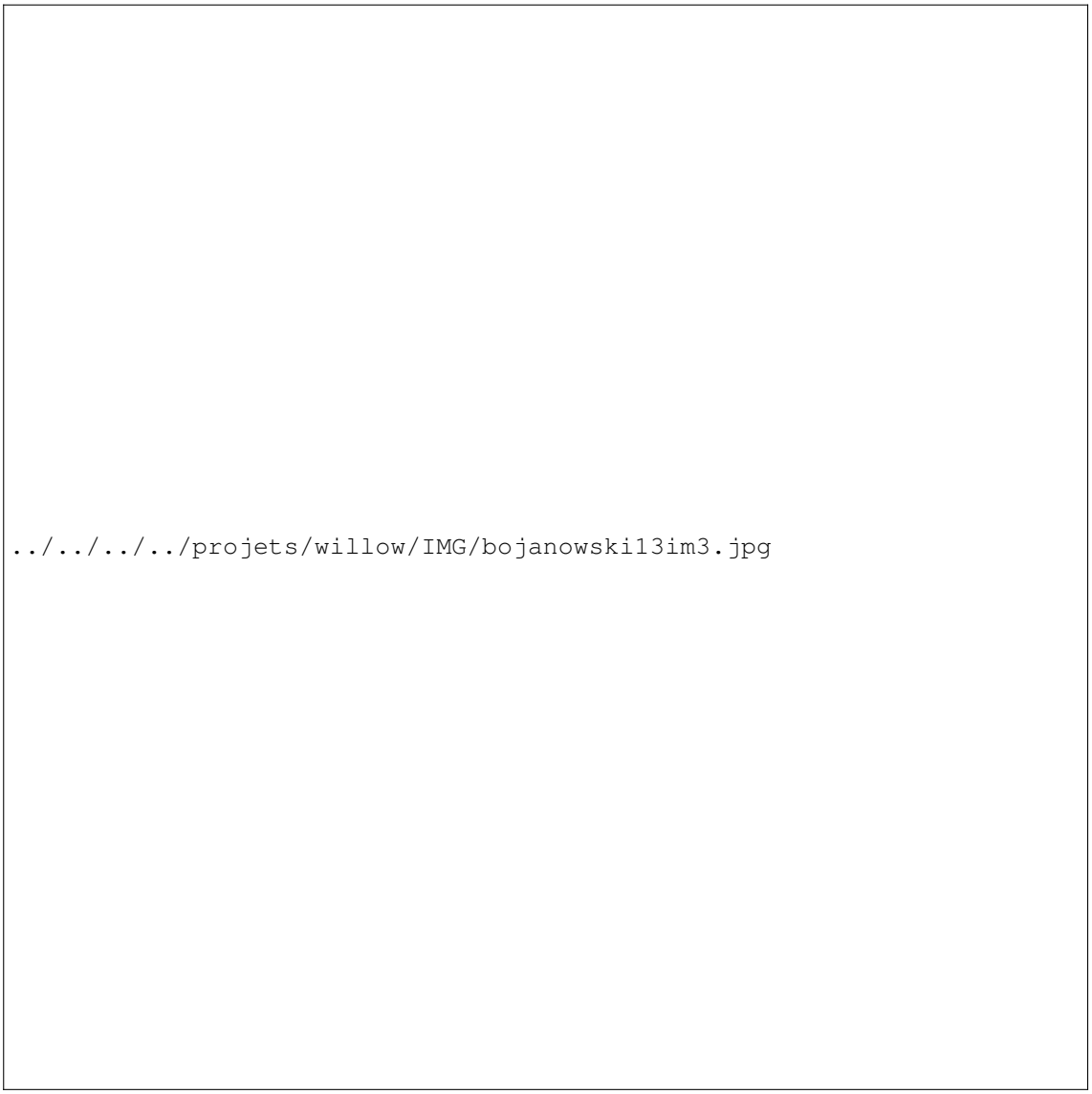


Figure 6. Examples of automatically assigned names and actions in the movie Casablanca. Top row: Correct name and action assignments for tracks that have an actor/action constraint in the script. Bottom row: Correct name and action assignments for tracks that do not have a corresponding constraint in the script, but are still correctly classified. Note that even very infrequent characters are correctly classified (Annina and Yvonne). See more examples on the project web-page: http://www.di.ens.fr/willow/research/actoraction/